

RESULT 1
PTEI_HUMAN
ID PTEI_HUMAN STANDARD; PRT; 319 AA.
014734; 015261;
16-OCT-2001 (Rel. 40, Created)
01-MAR-2002 (Rel. 40, Last sequence update)
DE Peroxisomal acyl-coenzyme A thioesterase (update)
DE (Peroxisomal long-chain acyl-coA thioesterase 1) (EC 3.1.2.2)
DE acyl coA thioesterase) (thioesterase II) (hTE).
GN PTEI OR HNACTE.
OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX NCBI_TaxID=9606;
RB [1]
RX SEQUENCE FROM N.A.
RA Kawanabe H., Shiratori T., Shoji H., Miyatake S., Okazaki Y.,
RT "A novel, acyl-coA thioesterase enhances its enzymatic activity by
RL direct binding with HIV Nef.";
RN Biochem. Biophys. Res. Commun. 238:234-239(1997).
RP [2]
RX SEQUENCE FROM N.A.
RA Liu L.X., Margottin P., Legall S., Schwartz O., Sellig L., Benarous R.,
RX MEDLINE=9729085; PubMed=9153233;
RA Benichou S.;
RT "Binding of HIV-1 Nef to a novel thioesterase enzyme correlates with
RT Nef-mediated CD4 down-regulation.";
RN J. Biol. Chem. 272:13779-13785(1997).
RP [3]
RX SEQUENCE FROM N.A.
RA Jones J.M., Nau K., Geraghty M.T., Erdmann R., Gould S.J.;
RX MEDLINE=99194760; PubMed=10092594;
RA "Identification of peroxisomal acyl-CoA thioesterases in yeast and
RA humans.";
RN J. Biol. Chem. 274:9216-9223(1999).
RP [4]
RX SEQUENCE FROM N.A.
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.B.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,
RA Coulson A.G., Coville G.J., Deadman R., Dhani P., Dunn M.,
RA Ellington D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levasianno M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,

ALIGNMENTS

RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,
RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.W., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Sacco D., Smith M.L., Soderglund C., Steward C.A., Sulston J.E.,
RA Swaby M., Thomas A.C., Vaidya K., Wallis J.M., Wallis J.M., Thorpe A.,
RA Tracey A., Thomas A.C., Vaidya K., Wallis J.M., Wallis J.M., Thorpe A.,
RA Whitehead S.L., Whitaker P., Walley J.L., Williams D., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RT Nature 414:865-871(2001).
CC -1- FUNCTION: MAY PLAY A ROLE IN FATTY ACID OXIDATION RATHER THAN
CC FORMATION OF FATTY ACIDS. MAY MEDIATE NEF-INDUCED DOWN-REGULATION
CC OF CD4.
CC -1- CATALYTIC ACTIVITY: Palmitoyl-CoA + H(2)O -> CoA + palmitate.
CC -1- SUBUNIT: INTERACTS WITH HIV-1 NEF.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- SIMILARITY: BELONGS TO THE C/M/P THIOESTER HYDROLASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENEM: AF014404; AB071665; 1;
CC ENEM: AB071665; 1;
CC ENEM: AB071665; 1;
CC ENEM: AF124564; AB071665; 1;
CC ENEM: AL008726; CA015502; 1;
CC HSP: P23911; 1C8U;
CC InterPro: IPR003703; Acyl-CoA thio.
CC Pfam: PF02551; Acyl-CoA thio.
KW Hydrolase; Serine esterase; Peroxisome.
FT ACT_SITE 78 78 BY SIMILARITY.
FT ACT_SITE 232 232 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT SITE 317 319 LWR -> VRS (IN REF. 2).
FT CONFLICT 291 293 L -> R (IN REF. 2).
FT CONFLICT 319 319 L -> R (IN REF. 2).
SQ SEQUENCE 319 AA; 35914 MW; 8345C65E5ABF3326 CRC64;

Query Match 97.3%; Score 1621; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.1e-138;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

11 PGDLRSVLTVTNLEPLDEDFRGHRYWPAKLEFGQIVGQALVAAAKSVSDVHVH 70
19 PGDLRSVLTVTNLEPLDEDFRGHRYWPAKLEFGQIVGQALVAAAKSVSDVHVH 78
71 SLHCYVRAGDPLKPLVLYQVETRTGSGFSVRSVAVQHGKPFIFTCQASQQAQSPMOH 130
79 SLHCYVRAGDPLKPLVLYQVETRTGSGFSVRSVAVQHGKPFIFTCQASQQAQSPMOH 138
131 QFSNTPVPPPEELDCETLIDYLRDPNQRKRYPLALNRITAAQGEVPIEKPNPSPLSOL 190
139 QFSNTPVPPPEELDCETLIDYLRDPNQRKRYPLALNRITAAQGEVPIEKPNPSPLSOL 198
191 QRMPEQNFWRVARGITGEGDMKMCVAAVTSYAFSLGTALLPHOMQHKVHFVSLDHS 250
199 QRMPEQNFWRVARGITGEGDMKMCVAAVTSYAFSLGTALLPHOMQHKVHFVSLDHS 258
251 MWFHAPFRADNMALTECESPWAGGSGRLVHGRLMRPDQGVLANVCAQEGVIRKPVQVSESK 310
259 MWFHAPFRADNMALTECESPWAGGSGRLVHGRLMRPDQGVLANVCAQEGVIRKPVQVSESK 318
311 L 311
319 L,319



Result No.	Query			DB	ID	Description
	Score	Length	Match			
1	1621	97.3	319	1	JC5644	acyl-CoA thiolester
2	612	36.7	200	1	A13272	acyl-CoA thiolester
3	583.5	35.0	294	2	AC3912	acyl-CoA thiolester
4	583.5	35.0	316	2	PF9689	acyl-CoA thiolester
5	568	34.1	289	2	B83153	acyl-CoA thiolester
6	567	34.0	343	2	T29819	hypothetical prote
7	559	33.6	310	2	A82734	conserved hypotet
8	558	33.5	286	2	AE6381	acyl-CoA thiolester
9	553	33.2	286	1	DA4775	acyl-CoA thioleste
10	553	33.2	286	2	B90692	acyl-CoA thiolester
11	553	33.2	286	2	F85542	acyl-CoA thiolester
12	543	32.6	286	2	AD3248	acyl-CoA thiolester
13	536	32.2	286	2	AD0560	acyl-CoA thiolester
14	518.5	31.1	430	2	T25623	hypothetical prote
15	501.5	30.1	284	2	G87701	acyl-CoA thiolester
16	499.5	30.0	320	2	C86148	hypothetical prote
17	460	27.6	286	1	B64047	acyl-CoA thioleste
18	442.5	26.6	281	1	D70570	probable acyl-CoA
19	440.5	26.4	465	2	T15540	hypothetical prote
20	327.5	19.7	365	2	T27460	hypothetical prote
21	322	19.3	213	2	T01216	hypothetical prote
22	311.5	18.7	297	2	H87068	acyl-CoA thiolester
23	294.5	17.7	300	1	H70557	probable acyl-CoA
24	273	16.4	349	1	S52763	hypothetical prote
25	243.5	14.6	276	2	T15536	hypothetical prote
26	211.5	6.9	323	2	D87472	hypothetical prote
27	112.5	6.8	250	2	T15887	hypothetical prote
28	101	6.1	265	2	D83287	hypothetical prote
29	100.5	6.0	208	2	T25035	hypothetical prote

QY 311 L 311

QY 311 L 311

Db 319 L 319

RESULT 2

AL3272

acyl-CoA thioesterase II (EC 3.1.2.-) [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C:Accession: AL3272

C:Query Match 36.7%; Score 612; DB 2; Length 300;

Best Local Similarity 44.4%; Pred. No. 1.le-46;

Matches 135; Conservative 52; Mismatches 89; Indels 28; Gaps 9;

Qy 7 TAALPFGDLRSVLTVTNLNLEPDLDFGRHYWPAKRLFGGQIVGQALVAARKSVSD 66

Db 12 TAA-----MREL--STLDELMDLFRGSPQGVQWQVFGQVIGQALIAAQTVDPE 64

Qy 67 VVHSLHCYFVRAGDKPLVLYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPS 126

Db 65 RHVHSLHAYFVRPGDAIPIYVDRIIDGSSFSRVLAKHGKAFTLSASF-OTDEG 123

Qy 127 PMQHFSMPT-VPPPELDCETLIDQYL--RDNPLKRYPLALNRTAAQVEPIEIKPVN 183

Db 124 GLDHQWNPBGLPQVQDHDHKEKYLDMAPGYRKYW-----ERERPIEIKPVN 175

Qy 184 PSLSOLQRMPEKQMFWRARGYIGEGDMKHCVAAYISDYAFGLTALLPHQVHVF 243

Db 176 LTHYFSREKLEPVQVWVRKGLVDP--DRALQALILAYLSMDTLTSLPH--GRFIP 231

244 -----MVSLSHMKHFRADHMLYECESPMWAGSGRLVHGRNQGVLAIVCAQES 298

232 DRMQVSLDLMHFRPCRLDDMLLYTODAPASGARGFNRGALYTRQGVILASVAQES 291

Qy 299 VIRV 302

Db 292 LIRV 295

RESULT 3

AC2915

acyl-CoA thioesterase II [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2001 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AC2915

C:Query Match 35.0%; Score 583.5; DB 2; Length 316;

Best Local Similarity 42.3%; Pred. No. 3.9e-44;

Matches 121; Conservative 56; Mismatches 96; Indels 13; Gaps 6;

Qy 19 LVYTVNLNLEPDLDFGRHYWPAKRLFGGQIVGQALVAARKSVSDVHVSILHCYFVR 78

Db 36 LIAT-LDLERLEENLYRGSPQIGWQRVFGGVINQALINAAQTVDVDRVHSLHAYFMR 94

Qy 79 AGDKPLVLYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPSMWHQFSMPTVP 138

Db 95 FQDPLPIYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPSMWHQFSMPTVP 153

Qy 139 PPELDCETLIDQYL--RDNPLKRYPLALNRTAAQVEPIEIKPVNPSLSOLQRMPEKQ 198

Db 154 PPEKLMSEQQAFL-----AKAPASIRKYWSNKRPIETIRPSVSLTHYSKEKLEPOD 207

Qy 199 FFWVRARGYIGEGDMKHCVAAYISDYAFGLTALLPHQVHVFHMFVSLDHSMPHFA 255

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AC2915

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-294 <KUR>

A:Cross-references: GB:AE008688; PIDN: AAL43737.1; PID: g17741270; GSPDB: GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: LesB

A:Map position: circular chromosome

C:Superfamily: acyl-CoA thioesterase II

Query Match 35.0%; Score 583.5; DB 2; Length 294;

Best Local Similarity 42.3%; Pred. No. 3.6e-44;

Matches 121; Conservative 56; Mismatches 96; Indels 13; Gaps 6;

Qy 19 LVYTVNLNLEPDLDFGRHYWPAKRLFGGQIVGQALVAARKSVSDVHVSILHCYFVR 78

Db 14 LIAT-LDLERLEENLYRGSPQIGWQRVFGGVINQALINAAQTVDVDRVHSLHAYFMR 72

Qy 79 AGDKPLVLYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPSMWHQFSMPTVP 138

Db 73 FQDPLPIYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPSMWHQFSMPTVP 131

Qy 139 PPELDCETLIDQYL--RDNPLKRYPLALNRTAAQVEPIEIKPVNPSLSOLQRMPEKQ 198

Db 132 PPEKLMSEQQAFL-----AKAPASIRKYWSNKRPIETIRPSVSLTHYSKEKLEPOD 185

Qy 199 FFWVRARGYIGEGDMKHCVAAYISDYAFGLTALLPHQVHVFHMFVSLDHSMPHFA 255

Db 186 IWVRARGV-PADPRLQSAIILAYLSMDTLTSLYANGITTFDPDSIQ-VASLDHAMPHR 243

Qy 256 PFRADHMLYECESPMWAGSGRLVHGRNQGVLAIVCAQESGVR 301

Db 244 PCRLDDMLLYTODSPSARGGLTRGNITFTRQGVILASVAQESGLIR 289

RESULT 4

F97689

acyl-CoA thioesterase II (PA3942) [imported] - Agrobacterium tumefaciens (strain C58)

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: F97689

C:Query Match 35.0%; Score 583.5; DB 2; Length 316;

Best Local Similarity 42.3%; Pred. No. 3.9e-44;

Matches 121; Conservative 56; Mismatches 96; Indels 13; Gaps 6;

Qy 19 LVYTVNLNLEPDLDFGRHYWPAKRLFGGQIVGQALVAARKSVSDVHVSILHCYFVR 78

Db 36 LIAT-LDLERLEENLYRGSPQIGWQRVFGGVINQALINAAQTVDVDRVHSLHAYFMR 94

Qy 79 AGDKPLVLYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPSMWHQFSMPTVP 138

Db 95 FQDPLPIYQVETRTGSSFSVSKAVQHKPFICQASFOQAQPSMWHQFSMPTVP 153

Qy 139 PPELDCETLIDQYL--RDNPLKRYPLALNRTAAQVEPIEIKPVNPSLSOLQRMPEKQ 198

Db 154 PPEKLMSEQQAFL-----AKAPASIRKYWSNKRPIETIRPSVSLTHYSKEKLEPOD 207

Qy 199 FFWVRARGYIGEGDMKHCVAAYISDYAFGLTALLPHQVHVFHMFVSLDHSMPHFA 255

A:Title: The Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacter

A:Reference number: A97359; PMID:11743194

A:Accession: F97689

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-316 <KUR>

A:Cross-references: GB:AE007869; PIDN: AAK88471.1; PID: g15157976; GSPDB: GN00169

C:Genetics:

A:Gene: AGR_C_4997

A:Map position: circular chromosome

C:Superfamily: acyl-CoA thioesterase II

D8	IIVRAGEV - PADPLOSAILAYSDTWLITLSYACGTTTFDSIQ-VASLDHAMFHR
Q7	PFRADHMLYECSPPWSSRGVLGHGLRWDRGVLAITCAOEGVR
D9	PCRUDDTLTDPSPSAGSGILTCGNFTIETROGEVLVAOVAELGR

RESULT 5
B83153 acyl-CoA thioesterase II PA3942 [imported] - Pseudomonas aeruginosa (strain PA01)
B83153 C:Species: pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83153
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yau, Y.; Brody, L.L.; Coulter, K.R.; Kas, A.; Folger, K.R.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:2043737
A:Accession: B83153
A:Status: preliminary
A:Feature type: ONTO>
A:Feature key: tag
A:Cross-references: GB:AE004812; GB:AE004091; NID:g9950125; PIDN:RAG07329.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: tsbB; PA3942
C:Superfamily: acyl-CoA thiolesterase II

Query Match 34.1%; Score 568; DB 2; Length 289;
Best Local Similarity 42.1%; Pred. No. 8.4e-43;
Matches 123; Conservative 48; Mismatches 95; Indels 26; Gaps 7;

QY	20	VTTVLNLEPDLDEOLFGRRHYVYPAKALEFGOILVGOALVRAAKNSVDVHVSHLCYFVRA	75
Db	8	LVALLSEPIEENLFKMSQDIFGFGYLGQVGLQAALSSQATVEDDRHHVHSHGCFILRP	67
QY	80	GDPKLPVLYQVETRTGSGFSYRSYKAVOHGKPFICQASFOQAQSPMQHQSMPVYP	139
Db	68	GDASMPVYQVDRVGGSGFSKRVATLQKQPIFTUSSFSQIDEEG-FHHQITEMPEVG	126
QY	140	PEBLDCETL-----IDYLRDPNLQKRYPLALNRIAAQAEVPIETKPNWSPISQLQR	192
Db	127	PDNLPSELRLTRQRIEPIERIKLL--YYP-----KPIETRPQTHDPYDPQP	173
QY	193	MEPKQFWYRARGYIGEGDMKHCCVAAAYISDYAFGLTALLPHQ---WQKHVFMVSLDH	249
I	174	DEPVKYLMFNRADGKLPDVP-ALHRYLLATVASFGLTTSQMOPHAVSYMQ-KTMQVASLDH	231
QY	230	SMWFHAPFRADHNLVPECSPWAGSGRLVHGRILWRDQGYLVATCAQEGVIR	301
Db	232	ALMFHDLRADENLLVADISWAGNARGFVRSIFNQSGLVASSQGLIR	283

RESULT 6
T29819
hypothetical protein F25E2.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21
T29819
R:Accession
P: Loc Tn
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of *C. elegans* cosmid F25E2.
A:Reference number: 220691

A:Accession: 129819
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-343 <MIN>
A:Cross-references: EMBL:U50197; PDN:AAA91256.1; CESP:P25E2.3
A:Experimental source: strain Bristol N2
C:Genetics:

A;Gene: CESP:F25E2.3
A;Introns: 9/3; 48/2; 66/3; 96/1; 176/3; 245/1; 311/1
C;Superfamily: acyl-CoA thiolesterase II

Query Match 34.0%; Score 567; DB 2; Length 343;
Best Local Similarity 38.4%; Pred. No. 1.3e-42;
Matches 126; Conservative 52; Mismatches 104; Indels 46; Gaps 9;

[illegible]

7

A82734 conserved hypothetical protein XF1021 (imported) - *Xylella fastidiosa* (strain 945C)
 C:Species: *Xylella fastidiosa*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: A82734
 R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide S
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A:Reference number: A82515; MUID:20345717
 A:Note: for a complete list of authors see reference number A59328 below
 C:Accession: A82734
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-310 <SIM>
 A:Cross-references: GB:AE003939; GB:AE003849; NID:g9105949; PIDN:AAF83831.1; GSPDB:
 A:Experimental source: GB:Strain 945C
 R:Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga,
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, L.E.A.; Carraro, D.M.; Carr
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000

[illegible]

Query Match 33.6%; Score 559; DB 2; Length 310;
Best Local Similarity 39.5%; Pred. No. 5.8e-42;
Matches 117; Conservative 54; Mismatches 107. Indels

```

21  TTVNLNLEPDLDFRGRHYWPAKRLFGQGIYQCALVAAAKSVSDVHVLSLHYCYFRAG 80
Db  TMLSLRLELDNLFQGNRDI GTKYVFGGGIQAALAAAGNTIENSRTAHSLHAYFLRTG 85
Qy  81  DPKLPVLQVHERTGGSSFSVRSKAVOHKPKIFTCQASQQQAQSPNQHQFSMTPTPPP 140
Db  82  TLLVLLIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII 145
Qy  86  NIQPIIINWERTROGKFSURKVAIOHGVIFECTASQECEDG-AEHQINMPAVPPP 144
Db  87  TLLVLLIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII 149
Qy  141  EELLDCETLLDQVLBDPMOKRYPALNRIAGQVPIEIKPVNPSISQLQRMPEKQMEW 200
Db  145  EDIATPPPL-----PPBALPMPQLQMTWNLISQGFERRVPPNNTLDKHSAPQLW 198
Qy  201  VRAGYITGSDGMKCCVAAVYISDYAFGLGTALLPHONQHKVIF-----NVSLDSHWMPH 254
Db  199  LRNNAPLGD-RIELHQILYASDQLFETYS-----RHGINTHTPOVQMSALDHALWF 253
Qy  255  APPRADHMLVIECSFPRAGSGRLVHGRLWRQDGVLAITCAQEGVIRKPVSESK 310
Db  254  RPRIDDWLLYLTIESPTAQRGLARGQFTFRNGVLNANTAQEGMLRTHNTHTPTSK 309

RESULT 8
acyl-CoA thioesterase II (EC 3.1.2.-) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0381
R:Parkhill, J.; Wren, B. W.; Thomson, N. R.; Titball, R. W.; Holden, M. T. G.; Prentis,
deno-Tarraga, A. M.; Chillingworth, T.; Cronin, A.; Davies, R. M.; Davis, P.; Doug-
lance, M. L.; Rutherford, K.; Sklumsom, M.; Skelton, J.; Stevens, K.; Whitehead, S.; B
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0381
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92376.1; PID:g15981079; GSPDB:GN00175
C:Genetics:
A:Gene: tesB
C:Superfamily: acyl-CoA thioesterase II
C:Keywords: thiolester hydrolase

```

Query Match	33.5%	Score 558;	DB 2;	Length 286;
Best Local Similarity	42.2%	Pred. No. 6.4E-42;		
Matches 122;	Conservative 56;	Mismatches 87;	Indels 24;	Gaps 9;
23	VLNLEPLDDELFGRHHVWPAKRLTGCTGQGVGAALAAKSVSESDVHVSHLSHCYFVRAGDP	82		
11	LDLDEKLEFGIFRQSESDGLGKQVGGVQGVGAIIYAAKQTVPALRTVHSEHSEYELRFGDS	70		
83	KLPVLQVQERTGSSFSVRSKVAOHGKPIFGICQASFOQAQSPHQHFSMPYTPPPEE	142		
71	SKPIIVDVLTKDNGSFAARRVSAIQNGKPIFYMTASF-QSQEGSEHQNTMPDVPPEEG	129		
143	LDLDETLQDVL--LSDPNQKRYPLALNRIAAQGVPEITKPNV-PSPLSQLORMEQKMF	199		
130	LMSCTDIAQFSLHILPEKRVF-----IGDP--FEMKPKYFAPL-QGSVEEENRVY	180		
200	WYRAGYIGEGDMKMCVAAVTSDFATGFTALLPHQOHVHPK-----VSLDISMKF	253		
181	WFRANGKMPD-DLRVHQYLLGVASDFNLPTALQP-----HGIGFLPGRGMQIALDISMKF	235		
254	HAPFRADHMLYECESPMWAGSRGLVHGRMLRQDQGLVATCAOBGVIRV	302		
236	HRPRDLDMLYAVESTSASGARGVRCQIINREGVLVASTVQSGVIRL	284		

RESULT 9
D64775
acyl-CoA thiolesterase (EC 3.1.2.-) II - *Escherichia coli*

C:Species: *Escherichia coli*
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 31-Mar-2000
C:Accession: D64775; JH0411; F50216
R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64775
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: type: DNA
A:Molecule type: DNA
A:Residues: 1-286 <BLAT>
A:Cross-references: GB:AE000151; GB:U00096; NID:g17886649; PIDN:AACT3555.1; PID:g1-g1
A:Experimental source: strain K-12, substrain MG1655
J. Nagert, J.; Narasimhan, M.L.; De Vaux, L.; Cho, H.; Randhawa, Z.I.; Cronan Jr
J. Biol. Chem. 266, 11044-11050, 1991
A:Title: Cloning, sequencing, and characterization of *Escherichia coli* thioesterase
A:Reference number: JH0411; MUID:91250410
A:Accession: JH0411
A:Molecule type: DNA
A:Residues: 1-11, 'OALYAAKATV', 12-41, 52-286 <NAG1>
A:Cross-references: GB:M63308; NID:g147931
A:Note: the sequence in Genbank entry ECOTESB, release 109.0, (PID:g147932) does "F:2-286/product: thioesterase II #status experimental <NAT>
F:56/Active site: His #status experimental
C:Genes: tesB
C:Complex: homotetramer
C:Function:
A:Description: hydrolyzes a broad range of acyl-CoA thioesters
C:Superfamily: acyl-CoA thioesterase II
C:Keywords: homotetramer; thioester hydrolase
F:2-286/product: thioesterase II #status experimental <NAT>
F:56/Active site: His #status experimental

Query Match 33.2%; Score 553; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 1.8e-41;
Matches 125; Conservative 51; Mismatches 91; Indels 28; Gaps 10:

Oy 15 LRSVLVTVTLNLEPDELDFGRHYHWYPAKRLFGQIVGQALVAAKSVSDVHSHLC 74
Db 5 LKML--TLNLEKTEEGFGRQSDGLGQVGGVQALYAKETVPEERLVHSFHS 62
Oy 75 YFRAGDPKLPVLYQVTRTGGSSFSRSKAVQNGKPIFCQASQQAQSPMQHPSM 134
Db 63 YLRFQDSKAPITFDVTLRDGNSFSARVAALONGKPIFTMTAGF-QNPFAGFEHQMT 121
Oy 135 PTVPYPEELDCETLIDQ---YLSPNLOKRYPLALNRIAQVPEIEKPVN-PSPLSQL 190
Db 122 PSAPADG-LPSETQIAQLAHLLPPVLKQF-----ICDRPEVPRVPEHPLKLG- 171
Oy 191 QRMPEQWFWVARGYTGEDQMKCCVAAXTSDYAFGLTALLPHQWQHKVPM----- 244
Db 172 HVAEPHQVWIRANGSPVD-DLARIWYLLGYASDLNPLVALQP----HGIFEPGQIQI 226
Oy 245 VSLDSHMFHAPRADHMLYCEPSHAGSRLGVRGLRMDRGDGLVATCAQGVIR 301
Db 227 ATIDSHMFHPRPNELNELLYSVESTSASSARGFVRGEFFYTDQGVYASIVQGVWR 283

RESULT 10
B90692
acyl-CoA thioesterase II [imported] - *Escherichia coli* (strain O157:H7, substrain
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: B90692
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and
A:Reference number: A99629; MUID:211562231 PMID:115268796

A:Accession: B90692
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <HAY>
 A:Cross-references: GB:BA000007; PID:N:BA033929.1; PID:q1335963; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain R1MD 050952
 C:Genetics:
 A:Gene: rcs0506
 C:Superfamily: acyl-CoA thioesterase II

Query Match 33.2%; Score 553; DB 2; Length 286;
 Best Local Similarity 42.1%; Pred. No. 1.8e-41;
 Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;

QY 15 LRSVLVTVLNLEPDELPRGRHWTPAKRLFGQIVGQALVAAKSVSDVHVHSLHC 74
 Db 5 LKLL--TLNLEKIEGLRGOSDLGRVGGVGGVQALVAAKTVPEERLVHSHFS 62
 QY 75 YFVRAGDKPLVLYOVERTTSGSVSRVSKAVQHGKPIFCOASFQOAPSPMOHFSM 134
 Db 63 YFLRPDSKPIIYDVTLRDGNSSARVAAIQNGKPIFYMTASF-QAPEAGFEHQTM 121
 QY 135 PTVPPEELDCETLIDQ---YLDPNLOKRYPLALNRIAAQEVPIEKPVN-PSPLSQL 190
 Db 122 PSAPADPG-LFSETYQAGSLHLLPVLKQF-----ICDRPLEVREVFHPLKG- 171
 QY 191 QRPKQMFVTRANGYIOEGDKMKCCVAAVYISDYAFGLTALLPHQOHKVHFM----- 244
 Db 172 HVAEPHQRWIRANGSVDP-DLVHGYLLGASDLNLFVALOP-----HGIGFLPGIOI 226
 QY 245 VSLDSHMFHAPRADHMLVYECESPMWAGSGRGLVHGRWRQDGLAVTCAQEGVIR 301
 Db 227 ATIDHSMFHRPFLNENLLYSVESTSSASSANGFVRGFTTQDGLVASTVQEGVMR 283

RESULT 11

F85542 acyl-CoA thioesterase II [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: F85542
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A83480; MUID:21074935; PMID:11200551
 A:Accession: B90692
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <STO>
 A:Cross-references: GB:AE005174; NID:q12513317; PID:AA654802.1; GSPDB:GN00145; UNKP:205
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: tesB
 C:Superfamily: acyl-CoA thioesterase II

Query Match 33.2%; Score 553; DB 2; Length 286;
 Best Local Similarity 42.1%; Pred. No. 1.8e-41;
 Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;

QY 15 LRSVLVTVLNLEPDELPRGRHWTPAKRLFGQIVGQALVAAKSVSDVHVHSLHC 74
 Db 5 LKLL--TLNLEKIEGLRGOSDLGRVGGVGGVQALVAAKTVPEERLVHSHFS 62
 QY 75 YFVRAGDKPLVLYOVERTTSGSVSRVSKAVQHGKPIFCOASFQOAPSPMOHFSM 134
 Db 63 YFLRPDSKPIIYDVTLRDGNSSARVAAIQNGKPIFYMTASF-QAPEAGFEHQTM 121
 QY 135 PTVPPEELDCETLIDQ---YLDPNLOKRYPLALNRIAAQEVPIEKPVN-PSPLSQL 190
 Db 122 PSAPADPG-LFSETYQAGSLHLLPVLKQF-----ICDRPLEVREVFHPLKG- 171

QY 191 QRMPEKQMFVTRANGYIOEGDKMKCCVAAVYISDYAFGLTALLPHQOHKVHFM----- 244
 Db 172 HVAEPHQRWIRANGSVDP-DLVHGYLLGASDLNLFVALOP-----HGIGFLPGIOI 226
 QY 245 VSLDSHMFHAPRADHMLVYECESPMWAGSGRGLVHGRWRQDGLAVTCAQEGVIR 301
 Db 227 ATIDHSMFHRPFLNENLLYSVESTSSASSANGFVRGFTTQDGLVASTVQEGVMR 283

RESULT 12

A82248 acyl-CoA thioesterase II VC1063 [imported] - Vibrio cholerae (strain N16961 serogroup

C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: A82248
 R:Heideberg, J.F.; Eelsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller,
 L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: A82248

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <HET>

A:Cross-references: GB:AE004187; GB:AE003852; NID:g9655516; PID:AAF94222.1; GSPDB:G1

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

A:Genetics:

A:Gene: VC1063

A:Map position: 1

C:Superfamily: acyl-CoA thioesterase II

Query Match 32.6%; Score 543; DB 2; Length 286;
 Best Local Similarity 41.0%; Pred. No. 1.4e-40;
 Matches 118; Conservative 48; Mismatches 100; Indels 22; Gaps 7;

QY 22 TVLNLEPDELPRGRHWTPAKRLFGQIVGQALVAAKSVSDVHVHSLHCYFRAGD 81
 Db 10 SLLOLEKLEGLYRGASENLGLPOVGGVIGQALSAARTVDSRIVSFHSFLYPGD 69
 QY 82 KPLPVLVYVERTTSGSVSRVSKAVQHGKPIFCOASFQOAPSPMOHFSNTPVPPPE 141
 Db 70 PERPIIYDVNMLRGKSSYRVKATQNGRPIFYLTATYHGDNPG-FEHQITMPVVPGE 128
 QY 142 ELDCETLIDQYLRDNLQKRYPLALNRIAAQEVPIEKPVN-SPLSQLOEMPEKQMF 200
 Db 129 NFASESELAQ-----LAHFLPEKLOKACGEKPIEMRPVTVINPLKP-EKAEPKQYLW 181
 QY 201 VVRVYIGESDKMKCCVAAVYISDYAFGLTALLPH-----QWQKHVFHMFSLDSHMF 253
 Db 182 IRANGDMPP-DOLHLYLLGASDGLAVTALPHGVSLMTKPFQ-----VATIDHSIWF 235
 QY 254 HAPRADHMLVYECESPMWAGSGRGLVHGRWRQDGLAVTCAQEGVIR 301
 Db 236 HRPEKMDMDLLFAIESPTASNTRGLVRGEIYDROGNLIVATVQEGVMR 283

RESULT 13

AD0560 acyl-CoA thioesterase II [imported] - Salmonella enterica subsp. enterica serovar Typhi

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: This species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AD0560

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Chur

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fa

, S.; Moutle, S.; O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0560
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <PNA>
 C:Cross-references: GB:AL513382; PIDN:CAD08924.1; PID:gl6501736; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0508
 C:Superfamily: acyl-CoA thiolesterase II

Query Match 32.28; Score 536; DB 2; Length 286;
 Best Local Similarity 41.88; Pred. No. 5.7e-40;
 Matches 123; Conservative 49; Mismatches 88; Indels 34; Gaps 9;
 QY 22 TVLNLPLEDDELFGRHVWYPAKFLGQIVGQALVAAKSVSDVHSHLCYFYRAGD 81
 DB 10 TLNLNLEIEGLFRGSGDLGRQVGGVQVQALVAAKETVPEARLVHSHFSYFLRPGD 69
 QY 82 PKLPLYOVERTRTGSGFSVRSKAVOHGKPFICQASFOQAQSPMQHQSMTVPPE 141
 DB 70 SQRPIYDVEVLDRGNSASRVAATONGKPIFYMTASFQAPFG-FEHQKTMPTAVGPE 128
 QY 142 ELLDCETLIDQ---YLRDNLQKRYPLALNRIAAQVPIETKVPV---NP-----SPLSQ 189
 DB 129 G-LPSETEINQSLAHLPLIUEKF-----LCORPLEIRPVFEHNPGLKHVSAPVRQ 179
 QY 190 LQRMKPKQFWVRARYIGEGDMKMHCCVAAVYSDYAFGLTALLPH--OMQHKVHFMVSL 247
 DB 180 V-----MIRANGTVPD--DIRVHLYLGTASDLNPLVNPQPHGIGLEKGTQIAYI 229
 QY 248 DRSMWRPAPRADHMLYECSPWAGSGRLVGRWRCQGVLANVTCQAEQVIR 301
 DB 230 DRSMWRPAPRADHMLYECSPWAGSGRLVGRWRCQGVLANVTCQAEQVIR 283

RESULT 14

T25623
 Hypothetical protein C37H5.13 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 19-May-2000
 C:Accession: T25623
 R:Davidson, S.; Gillam, B.
 submitted to the EMBL data Library, February 1997
 A:Description: The sequence of C. elegans cosmid C37H5.
 A:Reference number: 220058
 A:Species: Caenorhabditis elegans
 A:Molecule type: DNA
 A:Residues: 1-430 <DNA>
 C:Cross-references: EMBL:U08315; PIDN:AA842376.1; GSPDB:GN00023; CESP:C37H5.13
 A:Experimental source: strain Bristol N2; clone C37H5
 C:Genetics:
 A:Gene: CESP:C37H5.13
 A:Map position: 5
 A:Introns: 14/3; 30/3; 80/3; 132/2; 222/3; 283/3; 352/2
 C:Superfamily: acyl-CoA thiolesterase II

Query Match 31.18; Score 518.5; DB 2; Length 430;
 Best Local Similarity 37.08; Pred. No. 3.4e-38;
 Matches 121; Conservative 53; Mismatches 116; Indels 37; Gaps 8;
 QY 14 DLRSVLVTVLNLPLEDDELFGRHVWYPAKFL---FGQIVGQALVAAKSVSDVHSH 70
 DB 111 DIRAGLTUTLNUERVNRLNLNHLKGRNSLPVYVGGVIGQALVAAKSVSDVHSH 170
 QY 71 SLHCYFVRAGDKPLVLYOVERTRTGSGFSVRSKAVOHGKPFICQASFOQAQSPMQH 130
 DB 171 SLHSYFVQSGWVERPILVQVDRIDGKSECTRLVAKALQDGEAIFTVQISFHRPEADSI 230
 QY 131 QFSMTVPVPEELDCETLIDQVLRDNLN-----OKRYPLALNRIAAQVPIET 179
 DB 231 QLPMPVAPAPDSLEDSTFTFERIKNNANIPAAAMTGMFKQKEIPPAFFRI-----FSF 284

QY 180 KPVNPSPLSQORMEP-----KOMFVVRARYIGEGDMKMHCCVAAVYSDYAF 228
 DB 285 RPYDIDSYLCLKKDDHDTAGHGHTPTQATRSYVWIKANENIGD--DPLRLAAAYVSDATMI 343
 QY 229 GTALLPHQOHKHYKHPM-VSLDHSMMFHAP-FRADHMLYECSPWAGSGRLVHGRWLRQ 286
 DB 344 ETALPHSKRGPIPSMALTLIDHSIMHTDNFRVDDMLLYENHSTIAGGSGSLLECKLMTR 403
 QY 287 DGVLANVTCQAEQVIRV---RPQVSESK 310
 DB 404 DGRVFTSTOEARLITAKNPNRNSPAK 430
 RESULT 15
 GB7701
 acyl-CoA thiolesterase II [Imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: GB7701
 R:Nietman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; N., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: GB7701
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-284 <SPO>
 C:Cross-references: GB:AE005673; NID:gl3425405; PIDN:AAK25611.1; GSPDB:GN00148
 A:Gene: CS3649
 C:Superfamily: acyl-CoA thiolesterase II

Query Match 30.18; Score 501.5; DB 2; Length 284;
 Best Local Similarity 41.58; Pred. No. 6.5e-37;
 Matches 118; Conservative 54; Mismatches 99; Indels 13; Gaps 9;
 QY 23 VLNLEPDELDFGRHRYWYPAKFLGQIVGQALVAAKSVSDVHSHLCYFYRAGDP 82
 DB 10 ILDLEPEVNLGRVSPNDGFPRIEGLVIAQALLAAKTVPDV-CHSUHAYFIRPGDV 68
 QY 83 KLPVLYOVERTRTGSGFSVRSKAVOHGKPFICQASFOQAQSPMQHQSMTVPPEE 142
 DB 69 TAPVLYEVRAROGGTTTTRVAAIQHGEIFNLAASFQTPEDG-FHQSENPASVOPES 127
 QY 143 LLDCETLIDQVLRDNLQKRYPLALNRIAAQVPIETKVPNP-SPLSQORMKPKQFW 201
 DB 128 L---PTEAD-FLR--SLGDQIHPRKVAIAERPRVDIRMIDPQNPAPVAKSGTKQV-NM 180
 QY 202 RARYIGEGDMKMHCCVAAVYSDYAFGLTALLPHQ--WQHKVHFVMSLDHSMMFHAPFRA 259
 DB 181 RAKAPLAGD-DVKMQAALVADSNAMFESALPHGLTWTPGQIAASLDHMMFHHPNF 239
 QY 260 DHMMLYECSPWAGSGRLVHGRWLRQDGVLANVTCQAEQVIRK 303
 DB 240 NQMTLFAQDSPASQGRGLVRGQNSODGKLLASVAQECMLMRV 283

Search completed: August 19, 2002, 13:19:52
 Job time: 78 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:19:34 ; Search time 13.48 Seconds
(without alignments)
893.307 Million cell updates/sec

Title: US-09-766-366-1

Perfect score: 1666

Sequence: 1 MGRVATRALPPGULRSVLV.....VTCAGRGVIRKPVQVSESLK 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

T: number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1621	97.3	319	1 PTE1_HUMAN	O14734 homo sapien
2	1414.5	84.9	320	1 PTE1_MOUSE	P58137 mus musculus
3	553	32.0	285	1 TESB_ECOLI	P23911 escherichia
4	460	27.6	286	1 TESB_HAEIN	P44498 haemophilus
5	273	16.4	349	1 PTE1_YEAST	P41903 saccharomyc
6	99.5	6.0	417	1 TE63_HSVGA	P13199 herpesvirus
7	93	5.6	354	1 Y101_NPVOP	P24653 oryza p seu
8	90	5.4	454	1 PYRC_METHH	O27199 methanobact
9	89.5	5.4	2483	1 COA2_YEAST	O00763 homo septen
10	87	5.2	1654	1 PCFB_HUMAN	P12351 saccharomyc
11	85.5	5.1	527	1 TKX_HUMAN	O94913 homo sapien
12	85.5	5.1	1070	1 PTK1_HUMAN	O42681 homo sapien
13	85	5.1	1068	1 HRP2_HUMAN	O75146 homo sapien
14	84	5.0	648	1 CHA4_YEAST	P43634 saccharomyc
15	84	5.0	657	1 TRP_SCHPO	O13831 schizosacch
16	83.5	5.0	529	1 UD16_RAT	P08430 rattus norv
17	83	5.0	816	1 QA1F_NDUCR	P11638 neurospora
18	82.5	5.0	4466	1 DYRC_ANTIQR	P39057 antiochidari
19	82	4.9	1585	1 PK33_DICDI	P54675 dictyosteli
20	81.5	4.9	1177	1 MFD_BACSU	P37474 bacillus su
21	81	4.8	541	1 CREM_CANFA	P79145 canis famli
22	80.5	4.8	541	1 SD1L_PSEAE	O94uW5 pseudomonas
23	80	4.8	1581	1 MDP1_MOUSE	O93kY5 mus musculu
24	79.5	4.8	3591	1 WTS1_MOUSE	O93z24 mus musculu
25	79.5	4.8	3591	1 WTS1_MOUSE	O01886 cochlilobu
26	79.5	4.8	5217	1 GDB1_HUMAN	P08079 triticum ae
27	79	4.7	251	1 GDB1_HUMAN	O94v98 deinoceocccu
28	79	4.7	316	1 TKBE_HUMAN	O54108 mus musculu
29	79	4.7	364	1 TKBE_HUMAN	O54108 mus musculu
30	79	4.7	439	1 GFO_ZYMW6	O54108 mus musculu
31	79	4.7	860	1 WTS2_FANSD	P57062 zymomonas m
32	78.5	4.7	1350	1 NOS_DROME	O25771 drosophila
33	78	4.7	229	1 VG01_BPML5	O05218 mycobacteri

RESULT 1

TE1_HUMAN	TE1_HUMAN	STANDARD:	PRT:	319 AA.
AC	O14734	005261.		
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	03-MAR-2002	(Rel. 41, Last annotation update)		
DE	Peroxisomal acyl-coenzyme A thioester hydrolase 1 (EC 3.1.2.2)			
DE	Peroxisomal long-chain acyl-CoA thioesterase 1 (HIV-Nef associated)			
DE	acyl-CoA thioesterase (thioesterase II) (HFE).			
GN	PTE1 OR HNAACG.			
OS	Homo sapiens (human).			
OC	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97445158; Pubmed=9299485;			
RA	Watanabe H., Shiratori T., Shoji H., Miyatake S., Okazaki Y.,			
RA	Ikuta K., Sato T., Saito T.;			
RT	"A novel acyl-CoA thioesterase enhances its enzymatic activity by			
RT	direct binding with HIV Nef.";			
RL	Biochem. Biophys. Res. Commun. 238:234-239(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphoid;			
RX	MEDLINE=97298085; Pubmed=9153233;			
RA	Liu L.X., Margottin F., LeGall S., Schwartz O., Selig L., Benarous R.,			
RA	Benichou S.;			
RT	"Binding of HIV-1 Nef to a novel thioesterase enzyme correlates with			
RT	Nef-mediated CD4 down-regulation.";			
RL	J. Biol. Chem. 272:13779-13785(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Muscle;			
RX	MEDLINE=99194760; Pubmed=10092594;			
RA	Jones J.M., Nau K., Geraghty M.T., Erdmann R., Gould S.J.;			
RT	"Identification of peroxisomal acyl-CoA thioesterases in yeast and			
RT	humans.";			
RL	J. Biol. Chem. 274:9216-9223(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,			
RA	Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.G., Connor R., Corby N.R.,			
RA	Coulson A., Coville G.J., Deaman R., Dhani P., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Graham D.V., Griffiths M.N.D., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lenhavialho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,			

ALIGNMENTS

RESULT	1	78	4.7	317	1 YVCK_BACSU	006974 bacillus su
TE1_HUMAN	TE1_HUMAN	TE1_HUMAN	STANDARD:	PRT:	319 AA.	
AC	O14734	005261.				
DT	16-OCT-2001	(Rel. 40, Created)				
DT	16-OCT-2001	(Rel. 40, Last sequence update)				
DT	03-MAR-2002	(Rel. 41, Last annotation update)				
DE	Peroxisomal acyl-coenzyme A thioester hydrolase 1 (EC 3.1.2.2)					
DE	Peroxisomal long-chain acyl-CoA thioesterase 1 (HIV-Nef associated)					
DE	acyl-CoA thioesterase (thioesterase II) (HFE).					
GN	PTE1 OR HNAACG.					
OS	Homo sapiens (human).					
OC	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=97445158; Pubmed=9299485;					
RA	Watanabe H., Shiratori T., Shoji H., Miyatake S., Okazaki Y.,					
RA	Ikuta K., Sato T., Saito T.;					
RT	"A novel acyl-CoA thioesterase enhances its enzymatic activity by					
RT	direct binding with HIV Nef.";					
RL	Biochem. Biophys. Res. Commun. 238:234-239(1997).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Lymphoid;					
RX	MEDLINE=97298085; Pubmed=9153233;					
RA	Liu L.X., Margottin F., LeGall S., Schwartz O., Selig L., Benarous R.,					
RA	Benichou S.;					
RT	"Binding of HIV-1 Nef to a novel thioesterase enzyme correlates with					
RT	Nef-mediated CD4 down-regulation.";					
RL	J. Biol. Chem. 272:13779-13785(1997).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Muscle;					
RX	MEDLINE=99194760; Pubmed=10092594;					
RA	Jones J.M., Nau K., Geraghty M.T., Erdmann R., Gould S.J.;					
RT	"Identification of peroxisomal acyl-CoA thioesterases in yeast and					
RT	humans.";					
RL	J. Biol. Chem. 274:9216-9223(1999).					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,					
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,					
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,					
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,					
RA	Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,					
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,					
RA	Clegg S., Cobley V.E., Collier R.G., Connor R., Corby N.R.,					
RA	Coulson A., Coville G.J., Deaman R., Dhani P., Dunn M.,					
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,					
RA	Graham D.V., Griffiths M.N.D., Griffiths M.N.D., Gwilliam R., Hall R.E.,					
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,					
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,					
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,					
RA	Lenhavialho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,					

11	PGDLRSVLVTTLNLPDLDETLFRGRHYWPAKRLFCGOITVGOALVAAKKSVSDEVIHV	70
19	PGDLRSVLVTTLNLPDLDETLFRGRHYWPAKRLFCGOITVGOALVAAKKSVSDEVIHV	78
71	SLRUCFYVRAGDKPLVLYQVVERTTGGSSFSVRSAVQHGKPIIFICQASFOQAQSPMOH	130
79	SLRUCFYVRAGDKPLVLYQVVERTTGGSSFSVRSAVQHGKPIIFICQASFOQAQSPMOH	138
131	QFSMPTVPPPEELLDCETLIDXYLRDPLMKRLKRYPLALNRTAAQVEPTEIKVPNPSPLSQ	190
139	QFSMPTVPPPEELLDCETLIDXYLRDPLMKRLKRYPLALNRTAAQVEPTEIKVPNPSPLSQ	198
191	QRMPEKQNFWRARGYIGBGMKMHCCVAAAYISDYAFITGLTALLPHQWQHKVHPMVSLOHS	250
199	QRMPEKQNFWRARGYIGBGMKMHCCVAAAYISDYAFITGLTALLPHQWQHKVHPMVSLOHS	258
251	MNFHAPRADHMLYECESPWAGSGRGLVHGRLRQDQGLVATVCAQEGVTRVKPQVSGSK	310
259	MNFHAPRADHMLYECESPWAGSGRGLVHGRLRQDQGLVATVCAQEGVTRVKPQVSGSK	318

[illegible]

RESULT	3
TESB_ECOLI	
ID	TESB_ECOLI
AC	P23911;
DT	01-MAR-1992 (Rel. 21, Created)
STANDARD;	PRT;
	285 AA.

[illegible]

```

RT Influenzae Rd.*;
RL Science 269:496-512(1995).
CC -!- FUNCTION: CAN HYDROLYZE A BROAD RANGE OF ACYL-COA THIOESTERS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- PHYSIC: HOMOTETRAMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE C/M/P THIOESTER HYDROLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32693; AAC21752.1;
CC HSSP; P23911; 1C8U.
CC TIGR; H0076;
CC InterPro: IPR003703; Acyl_CoA_thio.
CC Pfam: PF02551; Acyl_CoA_thio; 1.
CC Hydrolase; Complete proteome.
CC ACT_SITE 58 58 BY SIMILARITY.
CC ACT_SITE 204 204 BY SIMILARITY.
CC SEQUENCE 286 AA; 32406 MW; 529390BF189CD07 CRC64;
CC
CC Query Match 27.6%; Score 160; DB 1; Length 286;
CC Best Local Similarity 36.9%; 50; Mismatches 109; Indels 22; Gaps 6;
CC Matches 106; Conservative 50;
CC
CC QY 23 VINLEPLEDELFGRHYWPAKRLFGQIVGQALVAAKSVSEDEVHSLHCYVRAGDP 82
CC Db 11 LKLEKIDDLIFRGESODLFGQVFGQVVAQLSNMVAPEDRILHSHAYFLAPGDS 70
CC
CC QY 83 KLPVLYOVERTRTGSSFSVRSKAVOHGKPIFCQASFCQAOQSPMQHQSFWPTPEPE 142
CC Db 71 QYPIIYVETLREGRNSALCVKAIOHKNTCHVTASFQVPEKG-FEHQNTMPNGAPED 129
CC
CC QY 143 LDCETLIDQYLRDNPLOKRYPLALNRITAAQVEPIK-----PVNPSLSQLQMEPK 196
CC Db 130 FTDENVMLQK-----VACTLPEPLNFAERPEVTRTKYLNPNFNGT-----KLPAE 177
CC
CC QY 197 QMEVWRARNGYIGEDMKHCVCVAATSDYAFGLTALLPHQ--WQHVKHFMVSLDHSWPH 254
CC Db 178 QYSWFKTNGET-PLDIKIQOGLLAYPSDPHGILTAHPHKEGFLQKMKVATIDHSIWFH 236
CC
CC QY 255 AFPRADHNMWLYCEESPWAGSGRLVHGRWRQDGLVAVTCAQEGVIR 301
CC Db 237 RPEDLNHHMLHATESNNAFGGRLAQGQIFSQDQLTATYQQEGLTR 283
CC
CC RESULT 5
CC PTBL_YEAST
CC ID PTBL_YEAST STANDARD; PRT; 349 AA.
CC AC P41903;
CC DT 01-NOV-1995 (rel. 32, Created)
CC DT 01-NOV-1995 (rel. 32, Last sequence update)
CC DT 16-OCT-2001 (rel. 40, Last annotation update)
CC DE Peroxisomal acyl-coenzyme A thioester hydrolase 1 (EC 3.1.2.2)
CC DE (peroxisomal long-chain acyl-coA thioesterase 1).
CC GN TEST OR PTBL OR YJ019C OR J1456.
CC OS Saccharomyces cerevisiae (Baker's yeast).
CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CC NX NCBI_TaxID=4932;
CC
CC SEQUENCE FROM N.A.
CC STRAIN-DH484.
CC RA H. Stumpf G. Domdey H.
CC RL submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC RN (2).
CC RP SEQUENCE FROM N.A.
CC RS STRAIN-5288C / FY1679;

```

```

RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN (3).
CC SEQUENCE FROM N.A., AND CHARACTERIZATION.
CC RX MEDLINE-99194760; PubMed=10092594;
CC RA Jones J.M., Nau K., Geraghty M.T., Erdmann R., Gould S.J.;
CC RT Identification of peroxisomal acyl-CoA thioesterases in yeast and
CC RT humans.;
CC RL J. Biol. Chem. 274:9216-9223(1999).
CC CC -!- FUNCTION: MAY PLAY A ROLE IN FATTY ACID OXIDATION RATHER THAN
CC CC FORMATION OF FATTY ACIDS.
CC CC -!- CATALYTIC ACTIVITY: Palmitoyl-CoA + H(2)O = CoA + palmitate.
CC CC -!- SUBCELLULAR LOCATION: Peroxisomal.
CC CC -!- SIMILARITY: BELONGS TO THE C/M/P THIOESTER HYDROLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X65972; CAAG0940.1;
CC EMBL; X67617; CAAG0940.1;
CC EMBL; Z49519; CAAG9543.1;
CC EMBL; A4124265; AAD27617.1;
CC SCOP; S0003780; TEST.
CC Pfam: IPR003703; Acyl_CoA_thio.
CC ID: PF02551; Acyl_CoA_thio; 1.
CC Hydrolase; Serine esterase; Peroxisome.
CC BY SIMILARITY.
CC ACT_SITE 347 349 MICROBODY TARGETING SIGNAL (POTENTIAL).
CC FT SITE 349 AA; 40259 MW; F1B5A51C9A46783E CRC64;
CC SEQUENCE 349 AA; 40259 MW; F1B5A51C9A46783E CRC64;
CC
CC Query Match 16.4%; Score 273; DB 1; Length 349;
CC Best Local Similarity 26.2%; Pred. No. 3.3e-17;
CC Matches 86; Conservative 53; Mismatches 135; Indels 54; Gaps 9;
CC
CC QY 23 VINLEPLEDELFGRHYWPA-----KRLFGQIVGQALVAAKSVSEDEVHSLHCYFV 77
CC Db 14 ILEVLPLSPTSFTVK--YLPAPVPGSKGFTGTLVSOSLLASLHTVPLNFFPTSLSHYFI 71
CC
CC QY 78 RAGDPLPVLVOVERTRTGSSFSVRSKAVOHGKPIFCQASF---QOAPSPMQHQSFM 134
CC Db 72 KGGDPRTKITYVONLNRGNFIHKQSAIYQHDKLFTSMILFAVQRSKEDHSLOHWETI 131
CC
CC QY 135 PTV-----PPEELDDCETLIDQYLRDNPLOKRYPLALNRITAAQVEPIETKVPNPSLSQL 190
CC Db 132 PGLOGKQDPHRYEATSLFQKLEVDLPQKLSRYASLSDRP--QOATSNKYVDAPFQYGV 189
CC
CC QY 191 ORMEPKOMFW-----VIRAGYI-----GEGDKM 214
CC Db 190 EYQPKMDPYSARHTDELDFYFKVRPITTVEHAGDESSLHKHHPYIPKSTPENARY 249
CC
CC QY 215 HCCVAAVYISDYAFGLTALLPHQOHKVH-FWVSLDHSMMFHAFPRADHNMWLYCEESPWAG 273
CC Db 270 NVAPATLSGSYLLCTPFRMLPLTCHFSVSLDHTITFRQLPHVANWLYLKISNPRSH 309
CC
CC QY 274 GSRLVHGLM-RQDGLVAVTCAQEGVI 300
CC Db 310 WDRHLVQGRKYFDTSGRINWASVQEGIV 337
CC
CC RESULT 6
CC TE63_HSVSA
CC ID TE63_HSVSA STANDARD; PRT; 417 AA.
CC AC P13199;
CC DT 01-JAN-1990 (rel. 13, Created)
CC DT 01-APR-1993 (rel. 25, Last sequence update)
CC DT 01-APR-1993 (rel. 25, Last annotation update)
CC DE 52 kDa immediate-early phosphoprotein.

```

57 OR EPRFL
 GS Herpesvirus saimiri (strain 11)
 NC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammarherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID:10363;
 [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE-92331688; PubMed-1321287;
 RA Albrecht J.-C.; Nicholas J., Biller D., Cameron K.R., Biesinger B.,
 RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
 RA Honess R.W.;
 RT "Primary structure of the herpesvirus saimiri genome.";
 RL J. Virol. 66:5047-5058(1992).
 [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE-92230228; PubMed-1314457;
 RA Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;
 RT "Analysis of nucleotide sequence of the rightmost 43 kbp of
 RT herpesvirus saimiri (HVS) c-DNA: general conservation of genetic
 RT organization between HVS and Epstein-Barr virus.";
 RL Virology 188:296-310(1992).
 [3]
 RP SEQUENCE OF 34-417 FROM N.A.
 RP MEDLINE-88300875; PubMed-2841477;
 RA Nicholas J., Gompels U.A., Craxton M.A., Honess R.W.;
 RT "Conservation of sequence and function between the product of the 52-
 RT kilodalton immediate-early gene of herpesvirus saimiri and the EBMLF1-
 RT encoded transcriptional effector (EB2) of Epstein-Barr virus.";
 RL J. Virol. 62:3250-3257(1988).
 -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
 HSV-2 UL54, EBV-1 5, VZV 4, EBV BMFL1, HCMV UL59, AND HSV-1 57.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 ENBL: X64346; CAA45680.1;
 DR ENBL: X86409; AAA46125.1; ALT_INIT.
 DR ENBL: X21943; AAA66558.1;
 PR: A36812; WMBEHA.
 KW Early protein; Transcription regulation; Phosphorylation.
 SQ SEQUENCE 417 AA; 46815 MW; 12FOEA3733PO0940 CRC64;

Query Match	6.0%; Score 99.5; DB 1; Length 417;
Local Similarity	23.9%; Pred. No. 0.16;
Matches	55; Conservative 30; Mismatches 82; Indels 63; Gaps 12;
Qy	91 ERTGTGSFVSRAVHOKPIFTCOASTQQAQSPSQMHQFS-----MFTVPPP 140
Db	91 KPRISNSTSRKTSNWR---VREAAQRRSPRFPKPSYHPRNGPLRNGPPRAPP 147
Qy	141 EELLDCETLIDQ-----YLDPNL-----QKRYALNRTIAAQEVPYIEK----- 180
Db	148 LKLFIDISILPKSGEPKFLPVLVPLSPCEAKNTKDYILVLAORAMHDVPYISSQLTANLL 207
Qy	181 PYNPSPLSQLQRMPEKQMFYRAR-GVIGEGDKMHCCVAAI-----SDYAF 227
Db	208 PVKFKPLSLISVRYTNYVYVSNRKETIASNL---CTVAALFDESICWGQQVLYKNDFF 264
Qy	228 -----LGTALLPHQHKRVHFVMSLDLSMWFHAPPRADHMLYEC 267
Db	265 SENGKDIILOETSALL-SOLVHKIK-MLPCKHCLMOTTP--QDHVTKQVC 310

RESULT	7
Y101_NPVOP	
ID	Y101_NPVOP
AC	P24653;
	STANDARD;
	PRT; 354 AA.

```

DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothesized 41.5 kDa protein in P6.5-Vp48 intergenic region (P40)
DE (ORF3) (ORF102).
DS Orygia pseudotsugata multicapsid polyhedrosis virus (OpNPV).
OS Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
ON NCBI_TaxID=164623;
RX [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=90188300; PubMed=2179466;
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohmann G.F.;
RA "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RA polyhedrosis virus genome.";
RT Virology 229:381-399(1997).
RL [2]
RL -!- SIMILARITY: TO CORRESPONDING ORF IN ACNPV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13959; BAA03056.1; -
DR EMBL: J47520; AAC52101.1; -
DR PIR: C34526; C34526.
DR Hypothetical protein. Late protein.
SQ SEQUENCE 334 AA, 39833 MW, 5606BDC31F73A39D CRC64;
Query Match 5.68; Score 93; DB 1; Length 354;
Best Local Similarity 21.98; P6; Mismatches 91; Indels 94; Gaps 12;
Matches 59; Conservative
QY 42 PAKRLGQIGVGAALVAANKSVSDYVHSHLCYFVRAGDPKLPVLVQVETRTGSSFSV 101
Db 90 PAQTVFG-----ALVSAPSAEPALYDMRR-----YRAAARRL-IQYSLNTTTSSEKVV 137
QY 102 RSV-----KAVCHKPIFICAOFSFOAOPSPMOH-----FSMPTVPP 140
Db 138 RDVVMTMTFLQRSENYHQLFKLLDTAMDFTCRPOLTEAQVSTLLHTRLTLEMTPTPD 197
QY 141 EEILD-----C-----ETLIDQVL--RDPNLOKRYPLAL 167
Db 198 MTTVDVRRSSFAFCFASPVLYRAKVVLQGETVSRDERTTLELLVERGDNIKLOPOQY 257
QY 168 NRLLAAQEVPI-----EIKPVNPSPLSOLQRMPEKQMFWRAGYI----- 207
Db 258 -VASGSEIIPDFCDPEFINRLKHLDPVLSHMYNAANSMEYTTMENAVANCKFNIEDY 316
QY 208 -----GEGDMKMHCCVAAVYISDY--AFLGT 230
Db 317 NRIFKGAESIKKANKATKADSDELDIYLG 346
RESULT 8
PYRC_METH
ID PYRC_METH STANDARD: PRT: 454 AA.
AC AC 027199; Q50488;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)

```

16-OCT-2001 (Rel. 40, Last annotation update)
 DHYDROOROTASE (EC 3.5.2.3) (DHase).
 PYRC OR MTHL127.
 OS Methanobacterium thermoautotrophicum.
 OS Archaea; Euryarchaeota; Methanobacteriaceae;
 OC Methanothermobacter.
 NCBI_TaxID=145262;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison O., Hoang L., Kagle P., Lum W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwen N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RP "Complete genome sequence of Methanobacterium thermoautotrophicum
 J. delat: functional analysis and comparative genomics.";
 J. Bacteriol. 179:7135-7155(1997).
 [2]
 RP SEQUENCE OF 24-454 FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=95014084; PubMed=7929010;
 RA Pithon D., Sharma S., Reeve J.N.;
 RP Growth phase-dependent transcription of the genes that encode the
 RP two-component enzyme N reductase isoenzymes and N5-
 RP methyltransferase in Methanobacterium thermoautotrophicum
 RP Methanobacterium thermoautotrophicum delta H.;
 J. Bacteriol. 176:6384-6391(1994).
 RL J. Bacteriol. 176:6384-6391(1994).
 CC -!- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O -> N-carbamoyl-L-
 CC aspartate.
 CC -!- COFACTOR: THIS ENZYME TIGHTLY BINDS ONE ZINC ATOM PER CHAIN WHICH
 CC IS REQUIRED FOR THE CATALYTIC MECHANISM. IT ALSO BINDS WEAKLY TO
 CC TWO OTHERS ZINCS WHICH ARE NOT ESSENTIAL FOR ACTIVITY
 CC (BY SIMILARITY).
 CC -!- PATHWAY: THIRD STEP IN PYRIMIDINE BIOSYNTHESIS.
 CC -!- SIMILARITY: BELONGS TO THE DHASE FAMILY. SUBFAMILY 2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 EMBL; AE000882; AAB85616.1; -;
 EMBL; U09990; AAA73440.1; -;
 MEKOPS; M38.972; -;
 DR InterPro; IPR002195; Dihydroorotase.
 DR Pfam; PF00744; Dihydroorotase; 1.
 DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
 DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
 KW Pyrimidine biosynthesis; Hydroxylase; zinc; Complete proteome.
 FT METAL 78 78
 FT METAL 80 80 ZINC (POTENTIAL).
 FT METAL 24 25 MF -> SL (IN REF. 2).
 FT CONFLICT 67 67 G -> A (IN REF. 2).
 FT CONFLICT 156 156 F -> N (IN REF. 2).
 FT CONFLICT 218 224 APPALAE -> CKTPRR (IN REF. 2).
 FT CONFLICT 263 266 CE -> SQ (IN REF. 2).
 FT CONFLICT 366 366 A -> P (IN REF. 2).
 FT CONFLICT 418 427 GPMVTIVRGR -> ARDHCOGO (IN REF. 2).
 SQ SEQUENCE 434 AA; 50534 MW; 40DE9F07B9DBEF83 CRC64;
 Query Match 5.48; Score 90; DB 1; Length 454;
 Best Local Similarity 22.74; Pred. No 1.3;
 Matches 48; Conservative 22; Mismatches 69; Indels 72; Gaps 8;
 QY 61 KSVSDVHVSILHC-----YFVAGDPLVLYOVRTGSSFSRSVKA--VQHG 110

DB 181 KDLSAPVPV-TVHCENRDVVMKSKELKDRSDPSAYALARPLAEVLSVAEVLALSIHHE 239
 QY 111 KPFIQCAQSPQAFPMQHQFSMTPTPPPELLDCTELLIDVLRDPLNOKR----- 162
 DB 240 HPVHICHLSTVKA-----LQLVEPREYVTCVTPPHLLDLSGAPRRGTWKTN 289
 QY 163 -----YPLALNRI-----AAQEVPIEIKPNVSP----- 186
 DB 290 PPLRPSPRIYPEFLDRINICTDHAPHGTEEKRGKIWDAPPGINLEVLKILLTLVSK 349
 QY 187 -----LSQALQM---EPKQMFVWRARGYIGEG 210
 DB 350 GRMSLSTIRRLAEPARIFGLRSGKRIAG 380
 RESULT 9
 COA2_HUMAN
 ID COA2_HUMAN STANDARD; PRT; 2483 AA.
 AC 000763; Q16852;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acetyl-CoA carboxylase 2 (EC 6.4.1.2) (ACC-beta) [Includes: Biotin
 DE carboxylase (EC 6.3.4.14)].
 GN ACACB OR ACC2 OR ACCB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX 10110
 RN 1;181_TaxID=9606;
 RS SEQUENCE FROM N.A.
 RC TISSUE=Adipose tissue; PubMed=9099716;
 RX MEDLINE=9756787; PubMed=9099716;
 RA Abu-Elheiga I., Almaraz-Ortega D.B., Baldini A., Wakil S.J.;
 RA Human acetyl-CoA carboxylase 2. Molecular cloning, characterization,
 RA chromosomal mapping, and evidence for two isoforms.;
 J. Biol. Chem. 272:10669-10677(1997).
 [2]
 RP SEQUENCE OF 1349-2134 FROM N.A.
 RC TISSUE=Adipose tissue; PubMed=8670171;
 RX MEDLINE=96765061; PubMed=8670171;
 RA Widmer J., Fassihi K.S., Schlichter S.C., Wheeler K.S., Crute B.E.,
 RA King N., Nutille-Mcmenemy N., Noll W.W., Daniel S., Ha J., Kim K.-H.,
 RA Witters L.A.;
 RA Identification of a second human acetyl-CoA carboxylase gene.*;
 Biochem. J. 316:915-922(1996).
 CC -!- FUNCTION: ACC-BETA MAY BE INVOLVED IN THE PROVISION OF MALONYL-COA
 CC OR IN THE REGULATION OF FATTY ACID OXIDATION, RATHER THAN FATTY
 CC ACID BIOSYNTHESIS. THIS PROTEIN CARRIES THREE FUNCTIONS: BIOTIN
 CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
 CC CARBOXYLTRANSFERASE.
 CC -!- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-)= ADP + phosphate
 CC + malonyl-CoA.
 CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
 CC -!- COFACTOR: BIOTIN.
 CC -!- PATHWAY: FIRST STEP (RATE LIMITING) IN LONG-CHAIN FATTY ACID
 CC SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; MAY ASSOCIATE WITH MEMBRANES.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE HEART, SKELETAL
 CC MUSCLES AND LIVER.
 CC -!- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).


```

CC RML: U9344; AAC58982.1;
DR RSP: P24182; 1DV1.
DR MIM: 601557;
DR InterPro: IPR001882; Biotin.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR000022; Carboxyl_trans.
DR Pfam: PF002765; Biotin_lipoyl.
DR Pfam: PF00364; Biotin_lipoyl.
DR Pfam: PF01039; Carboxyl_trans.
DR Pfam: PF00289; CPSase_L_chain.
DR Pfam: PF02786; CPSase_L_D2.
DR PROSITE: PS00188; BIOTIN.
DR PROSITE: PS00866; CPSASE.
DR PROSITE: PS00066; BIOTIN.
DR Fatty acid biosynthesis; Biotin: Alloative splicing;
KW ATP-binding; Phosphorylation; ATP (POTENTIAL).
FT ACT_SITE 458 463
FT BIND 584 584 BY SIMILARITY.
FT BINDING 927 927 BIOTIN.
FT DOMAIN 2095 2124 COENZYME A-BINDING (POTENTIAL).
FT MOD_RES 223 223 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 223 223 PHOSPHORYLATION (BY SIMILARITY).
FT VASPLIC 1115 1215 MISSING (IN SHORT ISOFORM).
FT CONFLICT 1365 1375 S -> N (IN REF. 2).
FT CONFLICT 1375 1375 T -> S (IN REF. 2).
FT CONFLICT 1408 1408 D -> G (IN REF. 2).
FT CONFLICT 1421 1421 P -> T (IN REF. 2).
FT CONFLICT 1450 1450 V -> M (IN REF. 2).
FT CONFLICT 1487 1487 P -> T (IN REF. 2).
FT CONFLICT 1487 1487 N -> S (IN REF. 2).
FT CONFLICT 1609 1609 N -> S (IN REF. 2).
FT CONFLICT 1630 1630 N -> F (IN REF. 2).
FT CONFLICT 1844 1844 M -> P (IN REF. 2).
FT CONFLICT 1917 1918 MI -> IM (IN REF. 2).
FT CONFLICT 1917 1918 MI -> IM (IN REF. 2).
SQ SEQUENCE 2483 AA; 279690 MW; A8C67827649E753 CRC64;

Query Match 54%; Score 90; DB 1; Length 2483;
Best Local Similarity 22.5%; Pos. No. 10;
Matches 62; Conservative 40; Mismatches 96; Indels 78; Gaps 15;
QY 22 TVLNFLELDELFRGRHYWPAKRLGGQIVGQALVAAKSVSEVHVHSLHCVFVRAGD 81
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 405 S-UTEMVEDTDLQGGKRSIVPEVDYDKG-----CVKVDGSLAAE-----RIGF 449
QY 82 PKLPIVLTQ-----VEKTRGSSFSV--RSVKAKQIKRPFICQASFOQAQSPMQHQS 133
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 450 PLMAKSSGGGKGIRESAEDPFLIFQVQSEIPGSPFLMKLA-----QH--- 497
QY 134 MPTVPPELLDCELTLDQY-----LRDFNQLKRYPLALNRTIAQAEVPLEKVPNPS 186
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 498 -----ARHLEVLQADQYGNVSLFGKDCISQRRH-----OKT-VEAPATPAT 540
QY 187 LSQLRMPKQFWFRAGYIGSDMKMKHCVAAYISDYAPLGLTALLPH-QNQRKHVHW 245
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 541 LAIFEMDQCAELAKTYOYVSAGTVEY-----LSQDGSFHFLELNPLOVEHCTEM 595
QY 246 S-----LDHSMKHFAPFR--ADHMLVRECSFW 271
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 596 ADVNLPAAQIQIAM--GAPLRLRLRLLYG-ESFW 628

RESULT 10
CYPL YEAST
ID CYPL YEAST STANDARD; PRT: 1483 AA.
AC P12351;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)

```

```

DE CYPL activatory protein.
GN CYPL OR HAP1 OR H1K25W.
OS Saccharomycetes; Baker's yeast.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RN MEDLINE=89125585; PubMed=2851659;
RN MEDLINE=91062241; PubMed=11024163;
RN PUBMED=11024163;
RN CYPL (HAP1) regulator of oxygen-dependent gene expression in yeast.
RT I. Overall organization of the protein sequence displays several
RT novel structural domains.
RL J. Mol. Biol. 204:263-276(1988).
RN [2]
RN SEQUENCE FROM N.A.
RN MEDLINE=91062241; PubMed=11024163;
RN PUBMED=11024163;
RN "Functional dissection and sequence of yeast HAP1 activator.
RL Cell 56:291-301(1988).
RN [3]
RN MUTANT CYPL-18.
RN MEDLINE=89125586; PubMed=2851659;
RN Verdier J., Galsane M., Guillard E., Defranoux N., Sionowski P.P.;
RN "CYPL (HAP1) regulator of oxygen-dependent gene expression in yeast.
RT I. Mutational activation of the gene and identification of
RT discriminating agents of gene control.
RL J. Mol. Biol. 204:277-282(1988).
RN [4]
RN DNA-BINDING.
RN MEDLINE=97042358; PubMed=8887558;
RN Zhang L., Guarente L.;
RN "The C6 Zinc Cluster dictates asymmetric binding by HAP1.
RT Structure.
RL J. Biol. Chem. 271:15467-15476(1996).
RN [5]
RN STRUCTURE BY NMR OF 60-100.
RN MEDLINE=96275662; PubMed=8683563;
RN Timmerman J., Vandeput A.-L., Bontems F., Lallemand J.-Y., Gervais M.,
RN Shechter E., Guillard B.;
RN "1H, 15N resonance assignment and three-dimensional structure of CYPL
RN (HAP1) DNA-binding domain.
RL J. Mol. Biol. 259:792-804(1996).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 55-135 OF MUTANT HAP1-18
RN COMPLEXED WITH DNA.
RN MEDLINE=99101378; PubMed=9886287;
RN King D.A., Zhang L., Guarente L., Marmorstein R.;
RN "Structure of HAP1-18-DNA implicates direct allosteric effect of
RN protein-DNA interactions on transcriptional activation.
RT. Struct. Biol. 6:22-27(1997).
RN [7]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 56-135 IN COMPLEX WITH DNA.
RN MEDLINE=99101385; PubMed=9886284;
RN King D.A., Zhang L., Guarente L., Marmorstein R.;
RN "Structure of a HAP1-DNA complex reveals dramatically asymmetric DNA
RN binding by a homodimeric protein.
RT. Struct. Biol. 6:64-71(1999).
RN [8]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 55-135 OF MUTANT HAP1-PC7
RN COMPLEXED WITH DNA.
RN MEDLINE=20482201; PubMed=11024163;
RN Lukens A.K., King D.A., Marmorstein R.;
RN "Structure of HAP1-PC7 bound to DNA: Implications for DNA recognition
RN and allosteric effects of DNA-binding on transcriptional activation.
RT Nucleic Acids Res. 28:3853-3863(2000).
RN [9]
RN FUNCTION: REGULATION OF OXYGEN DEPENDENT GENE EXPRESSION. IT
RN COMPLEXES WITH DNA AND IS A TRANSCRIPTION FACTOR.
RN THE SEQUENCE 5'-CGGNNNNKGG-3'.
CC 1- SUBUNIT: BINDS DNA AS A HOMODIMER.
CC 1- SUBCELLULAR LOCATION: Nuclear.
CC 1- MISCELLANEOUS: CYPL-18 MUTANT ACTIVATE THE EXPRESSION OF CYP3
CC (ISO-2) WHILE REDUCING THAT OF CYCL (ISO-1).
CC 1- MISCELLANEOUS: HEME IS AN EFFECTOR MOLECULE FOR CYPL/HAP1. THE

```



```

or send an email to license@isb-sib.ch
-----
CC CC
DR ENBL: L27071; AAB74557.1; -
DR ENBL: U34379; AAB60412.1; -
DR ENBL: U34367; AAB60412.1; JOINED.
DR ENBL: U34368; AAB60412.1; JOINED.
DR ENBL: U34369; AAB60412.1; JOINED.
DR ENBL: U34370; AAB60412.1; JOINED.
DR ENBL: U34371; AAB60412.1; JOINED.
DR ENBL: U34372; AAB60412.1; JOINED.
DR ENBL: U34373; AAB60412.1; JOINED.
DR ENBL: U34374; AAB60412.1; JOINED.
DR ENBL: U34375; AAB60412.1; JOINED.
DR ENBL: U34376; AAB60412.1; JOINED.
DR ENBL: U34377; AAB60412.1; JOINED.
DR ENBL: U34378; AAB60412.1; JOINED.
DR HSP: O06187; IAKX.
DR MM: 000058; -.
DR InterPro: IPR000719; Euk-pkinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001453; Tyr-kinase.
DR Pfam: PF000197; Pkinase; 1.
DR Pfam: PF00019; Pkinase; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00326; SH2; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
DR Transfaser: Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation.
KW
FT DOMAIN 14 19 POLY-CYS.
FT DOMAIN 68 73 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 82 142 SH3.
FT DOMAIN 150 246 SH2.
FT DOMAIN 271 527 PROTEIN KINASE.
FT NTRING 295 295 ATP (BY SIMILARITY).
FT NTRING 295 295 BY SIMILARITY.
FT ACT_SITE 390 390 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 420 420
FT CONFLICT 45 45 H -> R (IN REF. 2).
FT SEQUENCE 527 AA; 61239 MW; BC5D0705E4553003 CRC64;
Query Match 5.1%; Score 85.5; DB 1; Length 527;
Best Local Similarity 19.9%; Pred.No. 3.8;
Matches 62; Conservative 43; Mismatches 102; Indels 105; Gaps 16;
QY 89 QVETRTGSSFSVRS---VKAVQHKGKPIFTCAQSPQ-----QAQSPMHOFSMTVPVP 139
22 QKRWMTQISLTDELPERYQTHRRP-WLSQLSNKKQSTGRVQPSKRK---PLPLPVP 77
Db
QY 140 ---PEELLDCETLIDQYLRDP-----NLOKRYP-----LALNRP 170
78 SEVAEPKIQEAKYADYFLPREPCNLALRAEAYEYLIEKYNPHMKARDRLNGELIPSNV 137
QY 171 AAOEVP-TEFV-----KVPNPSPLOJQRMPEPKO-MFWVRAARYIGESDMKHCVCVAAYIS 223
138 TENKTKNLEIYEWYHNRITNQAEBHLRQSEKGAFIYRSHLGSTVISV-----188
QY 224 DYALGCT-----ALPLH-----QWQ-HKVFHWVSLDSKMTFH-----254
189 ---FWGARRSTAAAIKYQIKNDGOWMYAEHAFQISDELITYHQHNAAGLTRLRYP 245
255 -----APPDRAHMLVCEESPW-----AGCSGLGVGRLWRDQGLAVTCAQGV 299
Db 246 VGLMGSLCPATAGFSYKRWFDSPSELAFAIKETGSGRGVGVHLEWRSHTOVAIAWNGS 305

```

QY 300 TRVKPVSSEKL 311
 Db 306 MSEDFFIEAKV 317

RESULT 13
 PTK7 HUMAN
 AC PTK7 HUMAN STANDARD; PRT: 1070 AA.
 AC Q13308; O13417; PRT: 1070 AA.
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tyrosine-protein kinase-like 7 precursor (Colon carcinoma kinase-4)
 DE (CCk-4).
 DE PTK7 OR CCK4.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutharia; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 [11]

SEQUENCE FROM N.A.
 TISSUE=Colon carcinoma, and placenta;
 RX MEDLINE=96074849; PubMed=7478340;
 RA Mottis K., Jallat B., Alves F., Sures I., Plovman G.D., Ullrich A.;
 RT Colon carcinoma kinase-4 defines a new subclass of the receptor
 tyrosine kinase family;
 RN Oncogene 11:2179-2184(1995).
 [21]

SEQUENCE FROM N.A.
 TISSUE=Fibroblast;
 RX MEDLINE=97037064; PubMed=8882711;
 RA Park S.-K., Lee H.-S., Lee S.-T.;
 RT Characterization of the human full-length PTK7 cDNA encoding a
 receptor protein tyrosine kinase-like molecule closely related to
 chick Klg.
 RT
 RL J. Biochem. 119:235-239(1996).
 CC -!- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY
 THE CATALYTIC ACTIVITY OF TYROSINE KINASE. MAY BE CONNECTED TO THE
 PATHOPHYSIOLOGY OF COLON CARCINOMAS AND/OR MAY REPRESENT A TUMOR
 PROGRESSION MARKER.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, LIVER, PANCREAS,
 KIDNEY, PLACENTA AND MELANOCYTES. WEAKLY EXPRESSED IN THYROID
 GLAND, OVARY, BRAIN, HEART AND SKELETAL MUSCLE. ALSO EXPRESSED IN
 ERYTHROLEUKEMIA CELLS, BUT NOT EXPRESSED IN COLON.
 CC -!- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
 PROTEIN KINASES. TRK-TYPE SUBFAMILY.

THIS SWISS-PROT entry is copyright © 1996. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at European Bioinformatics Institute. There are no restrictions on use
 by non-profit institutions as long as its content is not used in a way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

EMBL: U33635; AAC87565.1; -;
 EMBL: U40271; AAC50484.1; -;
 HSSP: P12931; IPRM.
 MIM: 601890; -;
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR003600; Ig_c2.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Ig; 7.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM0410; Ig_c2like; 2.
 DR SMART: SM00108; Ig_c2; 5.
 DR SMART: SM00219; Tyrc; 1.
 DR SMART: SM00219; Tyrc; 1.

DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;
 Immunoglobulin domain; Repeat.
 FT CHAIN 1 30 POTENTIAL
 FT DOMAIN 31 1070 TYROSINE-PROTEIN KINASE-LIKE 7.
 FT DOMAIN 31 705 POTENTIAL
 FT DOMAIN 706 725 POTENTIAL
 FT DOMAIN 726 1070 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 146 108 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 143 208 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 239 308 IG-LIKE C2-TYPE DOMAIN 3.
 FT DOMAIN 336 398 IG-LIKE C2-TYPE DOMAIN 4.
 FT DOMAIN 426 488 IG-LIKE C2-TYPE DOMAIN 5.
 FT DOMAIN 517 577 IG-LIKE C2-TYPE DOMAIN 6.
 FT DOMAIN 606 671 IG-LIKE C2-TYPE DOMAIN 7.
 FT DOMAIN 736 1066 PROTEIN KINASE; INACTIVE.
 FT DISULFID 53 101 BY SIMILARITY.
 FT DISULFID 150 200 BY SIMILARITY.
 FT DISULFID 246 301 BY SIMILARITY.
 FT DISULFID 343 391 BY SIMILARITY.
 FT DISULFID 433 481 BY SIMILARITY.
 FT DISULFID 524 570 BY SIMILARITY.
 FT DISULFID 613 664 BY SIMILARITY.
 FT CARBOHYD 116 175 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 176 193 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 194 214 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 92 92 P -> R (IN REF. 2).
 FT CONFLICT 147 147 K -> T (IN REF. 2).
 FT CONFLICT 207 207 S -> G (IN REF. 2).
 FT CONFLICT 495 496 VL -> RV (IN REF. 2).
 FT CONFLICT 515 515 G -> E (IN REF. 2).
 FT CONFLICT 834 834 S -> T (IN REF. 2).
 FT CONFLICT 881 881 E -> P (IN REF. 2).
 FT CONFLICT 969 969 A -> G (IN REF. 2).
 FT CONFLICT 932 932 S -> F (IN REF. 2).
 SQ SEQUENCE 1070 AA; 118260 MW; 47CDF25B8E369BA5 CRC64;

Query Match 5.1%; Score 85.5; DB 1; Length 1070;
 Best Local Similarity 21.6%; Pred. No. 3.1; 105; Indels 119; Gaps 18;
 Matches 72; Conservative 37; Mismatches 37

QY 1 NCRVATAALPG-DLRSLVLTTLN-----LEPLDDELPRGRHWYPAKELFGQ 50
 Db 1 NGAAGSPARPRLLPLSLVLLPLGLGTQTATVIRKQPSQDALQGR----- 47
 QY 51 IGVGALVAAKSVSEVDVHSLHCYFVRAGDKLPVLVYQVTRTGTSSFSVRSYKAVOH- 109
 Db 48 ---RALLRCEVAPGPVH-----YMLDGS---APVODTERFAAGSSLSFAAVDPLQDS 96
 QY 110 -----GKPIFICQASFO-----QAQSPMQHQSMTPTVPPPEEL-LDCELTID- 151
 Db 97 GTFCVARDVDVTGEARSANASFNKIKWIEAGPVVILKHPASEAEIQPTQVTKURCH--IDG 154
 QY 152 -----QYLDRNPKRYP-----ALNRITAAQVEPIKIPVNPSPLSQLQRMPEKQMFV 201
 Db 155 HPRFTYQWFRDGT-----PLSDGDSNHTVSSKERNLTLRPAQ----- 192
 QY 202 RARGYTGEIDMMHICV-AA------ISDYAFGLGTALLPHQNHQKVPWVSLD 248
 Db 193 -----EHSGLYSCAHSAFSQACSSONFTLSTADESFARVVLAPOD-----VVYARY 239
 QY 249 HSMVFHAPFRAD-----HWMLYCESPWAGSR 276
 Db 240 BEAMFCQFSQAPPPSLQW-LFEDETFTINRSR 271

RESULT 14
ID HIPR_HUMAN STANDARD: PRT; 1068 AA.
AC 075146; Q9UED9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Huntingtin interacting protein 1 related (Hipl-related) (Hipl 12).
GN HPIR OR HPI2 OR KIA0655.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20515263; PubMed=11063258;
RA Chopra V.S., Metzler M., Rasper D.M., Engqvist-Goldstein A.E.Y.,
RA Singara R., Gan L., Fichter K.M., McCutcheon K., Dublin D.,
RA Nicholson D.W., Hayden M.R.;
RT HPI2 is a non-proapoptotic member of a gene family including HPI1,
RT an interacting protein with huntingtin.;
RL Mamm. Genome 11:1006-1015(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=brain;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kikuchi K., Ohta O.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.;
RL DNA Res. 5:169-176(1998).
RN [3]
RP SEQUENCE OF 179-1068 FROM N.A.
RX TISSUE=Neuroblastoma;
RA Seki N., Muramatsu M., Sugano S., Suzuki Y., Nakagawara A., Ohhira M.,
RA Hayashi A., Hori T., Saito T.;
RT "Cloning, expression analysis, and chromosomal localization of HPIR,
RT an isoform of huntingtin interacting protein (HPI1).";
RL J. Hum. Genet. 43:268-271(1998).
CC -1- FUNCTION: COMPONENT OF CLATHRIN-COATED PITS AND VESICLES, THAT MAY
CC LINK THE ENDOCYTIC MACHINERY TO THE ACTIN CYTOSKELETON.
CC -1- SUBUNIT: INTERACTS WITH ACTIN AND HUNTINGTIN INTERACTING PROTEIN 1
CC (HPI1). DOES NOT INTERACT WITH HUNTINGTIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. MEMBRANE-ASSOCIATED PROTEIN,
CC MAINLY LOCALIZED AT THE ENDOCYTIC COMPARTMENTS AND IN THE
CC PERINUCLEAR REGION (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: BRAIN, HEART, KIDNEY, PANCREAS, AND LIVER, BUT
CC NOT IN LUNG OR PLACENTA.
CC -1- DOMAIN: THE TALIN-LIKE DOMAIN BINDS F-ACTIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SLA2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AB014553; BAA31530.1; ALT_INIT.
CC EMBL; AB013384; BAA33713.1;
CC MIM; 603613;
CC InterPro; IPR001026; ENTH.
CC InterPro; IPR002558; ILWEQ.
CC Pfam; PF01417; ENTH; 1.
CC Pfam; PF01609; ILWEQ; 1.
CC ProDom; PD011820; ILWEQ; 1.
CC SMART; SM00273; ENTH; 1.
CC SMART; SM00307; ILWEQ; 1.

KW Actin-binding; Coiled coil.
FT DOMAIN 347 599 COILED COIL (POTENTIAL).
FT DOMAIN 800 1068 TALIN-LIKE.
SQ SEQUENCE 1068 AA; 119388 MW; 3CB7CF1191BFF8F CRC64;

Query Match 5.1%; Score 85; DB 1; Length 1068;
Best Local Similarity 19.0%; Pred. No. 10;
Matches 72; Conservative 47; Mismatches 123; Indels 136; Gaps 14;

QY 14 DLRSVLVTVTINLEPLD-----EDLFRGRHWYVPAKRLFGQQTIVGOALVAAAKSVSDV 67
Db DVANIFQLTFEMFDYMDCKLSEVFQRLNTALANSQMSQCLRAPIQLVIOQCSHLY 227
QY 68 H-----VSLHCIFVRAGDKPLVLY-----OVERTR 94
Db 228 HYTKLLKLLKLSCLPADTLOGRDRFHEQFSLNFRASD---MLFKELIQILPLP 283
QY 95 TGSFSFVSRYAVQVKQKPFIFC-QASFOQAQSPHQSFMTVPPEELLOCTLTDOY 153
Db 284 EGPNFLASALAEHLKPVVITPEAPEDEEF---ENLFIETGPPAGEVNVADLPDQT 340
QY 154 LRDPN-----LQ-----KYPDLALNRIAAQGVPIEKPV----- 182
Db 341 FGPNGSVKDDDDLDLIESLKREVMURSELEKIKLEAQRYIAQLKSQVNALRGLEPQRK 400
QY 183 -----NPSPLSOLQRMPEKQWVVRARYGIGEGDMKMHCCVAAVYISDYAFLGTALIP 234
Db 401 QKQKALVDNEQLRHLELAQLRAAQLEGRSOGRLREAPRKASATEARY----- 447
QY 235 HOWOHVHVEMVSLDSMTWHPFAPFRADHMLVECESPWAGSGRLVHGRLWRQDG----VL 290
Db 448 NKLKEK-----HSELVH-----VHAELLRKNADTAQOL 475

QY 291 AVT-CAQEGVIRVRPQVS 307
Db 476 TVTQSQSEEVARVKEQLA 493

RESULT 15
ID CHA4_YEAST STANDARD: PRT; 648 AA.
AC P43634;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CHA4 activatory protein.
GN CHA4 OR YLR098C OR L8004.8.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA Schjerling P.;
RN Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA Johnson W., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Hallsworth K., Fulton I., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Johnston L., Langston V., Hillier L., Jier M., Johnson D.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Taich A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RN Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hennes V., Rechmann S., Nentwich U., Schwager C., Ansorge W., Voss H.;
RN Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: ACTIVATES THE CHA1 GENE FOR L-SERINE DEHYDRATASE. BINDS

```

CC      TO THE DNA SEQUENCE 5'-GVCAGARAYRTRATTCRC-3'.
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC      CLUSTER DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-affiliated institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: Z49975; CAA90276.1; -
CC      DR      EMBL: U53876; AA067542.1; -
CC      DR      EMBL: Z73270; CAA97662.1; -
CC      DR      TRANSFAC: T02848; -
CC      DR      SGD: S0004088; CHA4.
CC      DR      InterPro: IPR001138; zn2_cy6_fungal.
CC      Pfam: PF00172; zn_clus; 1.
CC      SMART: SM00065; GAI4; 1.
CC      DR      PROSITE: PS00463; ZN2_Cy6_FUNGAL_1; 1.
CC      KW      PROSITE: PS00048; ZN2_Cy6_FUNGAL_2; 1.
CC      TR      Transcription regulation: Activator; DNA-binding; Nuclear protein;
CC      ZINC      Zinc; Metal-binding.
CC      FT      DOMAIN 32 35 POLY-ASN.
CC      FT      DNA_BIND 44 70 ZN(2)-CYS(6), FUNGAL-TYPE.
CC      FT      DOMAIN 48 52 POLY-ARG.
CC      FT      DOMAIN 105 109 POLY-SER.
CC      FT      DOMAIN 195 198 POLY-GLN.
CC      FT      DOMAIN 302 306 POLY-SER.
CC      SQ      SEQUENCE 648 AA; 74392 MW; CF3381B4C0356F82 CRC64;

```

Query Match
 Best Local Similarity 21.9%; Score 84; DB 1; Length 648;
 Matches 48; Conservative 32; Mismatches 49; Indels 90; Gaps 14;

```

QY 131 QFSNPTVPPPEELDCETLIDQYLRDPN--LQKRYPL---ALNRIAAOEVPETI---KP 181
Db 394 RLNSNTPVPTDELPEIETGIEIYDPKVIILSTANPLKRLIVLSRIT-----EIFASKI 447
QY 182 VNPSPUSQLQRME-----PKQMPVWRARGYIGEDMKHCCVAAAYISDY 225
Db 448 FSPNE-TLLQSEYLAQFNLEVYNWRDLPPQLQWK-----RSLMEM-----TDF 492
QY 226 AFLGTALLPHQWQKHVFWSLDSHSMWHPAFRADHMLYECESPMAGSGRLVHG----- 281
Db 493 ----NPTAYVWFHYIVLIS-----YKPF-----IYEIKQ-----SRELVEGYVDE 531
QY 282 -----FLNRQ-----DGLAVTCAQEGVIR 301
Db 532 LYLLKWKNKFTFEKATINYSAILAIOCKSNLIK 570

```

Search completed: August 19, 2002, 13:23:28
 Job time: 234 sec

Result No.	Query	Score	Match	Length	DB	ID	Description	SUMMARIES	
1	582	34.9	294	16	Q92L92	Q92L92	rhizobium		
2	568	34.1	289	16	Q9HX74	Q9HX74	pseudomonas		
3	567	34.0	343	5	Q19781	Q19781	caenorhabditis		
4	565	33.9	286	16	Q98B09	Q98B09	rhizobium		
5	562	33.7	290	2	Q9LO80	Q9LO80	streptomyces		
6	559	33.6	310	16	Q9PER7	Q9PER7	caenorhabditis		
7	543	32.6	286	16	Q9RT42	Q9RT42	vibrio cholerae		
8	541.5	32.5	488	2	Q9KX12	Q9KX12	streptomyces		
9	516.5	31.1	4	5	Q93058	Q93058	caenorhabditis		
10	506.5	30.6	168	16	Q92D86	Q92D86	caenorhabditis		
11	499.5	30.0	320	10	Q9LO87	Q9LO87	streptomyces		
12	463	27.8	292	16	Q9NM69	Q9NM69	pasturella		
13	449.5	27.0	368	3	Q9C321	Q9C321	cocciliobolus		
14	442.5	26.6	281	16	Q06209	Q06209	mycobacter		
15	440.5	26.4	346	5	Q9B1A5	Q9B1A5	caenorhabditis		
16	440.5	26.4	357	5	Q9B1A9	Q9B1A9	caenorhabditis		

Db 1 MSRPTEAT--PMD---ALLAILDLKLEENLFRGLSPQVQWRQGVGGQVIGQALVAAQ 54
 Qy 61 KVSSEVDVHSLHCYFVRAGDPKLPVLVQVVERTTSGSSFSVRKAVQHGKPIFCQASF 120
 Db 55 RYVDGGRYVHSLHAYFMRPGDPSVPIYEDVIRDSGFATRRVATQHGKAIFAMSASF 114
 Qy 121 QAAQSPMQHQFSMTVPVPEELLDCELTIDQYLRDNLQKRYPLALNRITAAQVEPIEK 180
 Db 115 QYDSDG-FRHFQDPDPVMPETLFGQELKEFL-----VHAPEAIRYWERPRPIEIR 167
 Qy 181 PVNPSLSOLQRMPEKMPFWVRARYIGBDMKHCVAAYISDYAFAGTALLPH--QK 237
 Db 169 PVLSLEYFRAKASPQDYYKAVGVTPD-EKHQAVALYLSDMTLLDTSLYANGTSYV 226
 Qy 238 QHKVHFWSLDRSMWFHAFRADHMLYECSPWAGSGRLVHGLRWODGVLAVTCAQE 297
 Db 227 DRSLSQ-VASLDHAMFHRSPKMDMLLYTQDPSAHGARGMTRGSLFDRSGVLITASVAGE 285
 Qy 298 GVIRVK 303
 Db 286 GLIRKK 291

RESULT 2

Q9HX74 Q9HX74 PRELIMINARY; PRT; 289 AA.
 ID Q9HX74
 AC Q9HX74
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ACYL-COA THIOESTERASE II.
 GN TESB OR PA3942.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brickman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen";
 RT Nature 406:959-964(2000).
 RM EMBL; A6004812; AAC07329.1; -.
 DR HSSP; P23911; 1C8U.
 DR InterPro; IPR003703; Acyl_CoA_thio.
 DR Pfam; PF02551; Acyl_CoA_thio; 1.
 KW Complete proteome.
 SQ SEQUENCE 289 AA; 32934 MW; 6592E1FFCE32DFE0 CRC64;

Query Match 34.18; Score 568; DB 16; Length 289;
 Best Local Similarity 42.18; Pred. No. 9.7e-49;
 Matches 123; Conservative 48; Mismatches 95; Indels 26; Gaps 7;

Qy 20 VTVVNLPLEDLDLFRGHVYVPAKRLFGGQVIGQALVAAKSVSEVDVHSLHCYFVRA 79
 Db 8 LVALLSLEPIENLRKMSODGFRGLVGLQVIGQALVAAKSVSEVDVHSLHCYFVRA 67
 Qy 80 GDEPLVLYQVVERTTSGSSFSVRKAVQHGKPIFCQASFQAAQSPMQHQFSMTVPV 139
 Db 68 GDASMPVYQVDRDGGSFSTRVATQKQPIFLSSQYDEG-FHQIEMPEVG 126
 Qy 140 PEELDCELT-----IDQYLRDNLQKRYPLALNRITAAQVEPIEKVNPSPLSQLR 192
 Db 127 PDLNPELELTQRAERIPERIRDKLL--YP-----KFIEIRVTDHPDYDPO 173

Qy 193 MEPKMPFWVRARYIGBDMKHCVAAYISDYAFAGTALLPHQ---WQHKVHFWSLDH 249
 Db 174 DEPKVYLFNRADGDLDPV-ALHKYLLAYASDFGLTTSQHPAVSVWQ-KTMQVASLDH 231
 Qy 250 SMWFHAFRADHMLYECSPWAGSGRLVHGLRWODGVLAVTCAQEGVIR 301
 Db 232 ALMFHRDLRADEMLYLAIDSPWAGNARGFVRGIFNQSGOLVASSQSGRLR 283
 RESULT 3
 Q19781 Q19781 PRELIMINARY; PRT; 343 AA.
 ID Q19781
 AC Q19781
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 38.8 KDA PROTEIN.
 GN F25E2.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=95069613; PubMed=9851916;
 RA None;
 RA "Genome sequence of the nematode C. elegans: a platform for
 investigating biology of the C. elegans Sequencing Consortium.";
 RT Science 284:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Leibach D. Minx M.;
 RT "The sequence of C. elegans cosmid F25E2.";
 RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL EMBL; U50197; AAA91256.1; -.
 DR HSSP; P23911; 1C8U.
 DR InterPro; IPR003703; Acyl_CoA_thio.
 DR Pfam; PF02551; Acyl_CoA_thio; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 343 AA; 38845 MW; 239065057C69E8A0 CRC64;

Query Match 34.08; Score 567; DB 5; Length 343;
 Best Local Similarity 38.48; Pred. No. 1.5e-48;
 Matches 126; Conservative 52; Mismatches 104; Indels 46; Gaps 9;

Qy 14 DLRSVLVTVVNLPLEDLDLF-----RGRHYVYVPAKRLFGGQVIGQALVAAKSVSEDV 67
 Db 27 DIRAGLIDTFLNQLQRIDTNLRIARHLKGRHSY---NAVYGGVYVQSLAAATAVDECF 83
 Qy 68 HVHSLHCYFVRAGDPKLPVLVQVVERTTSGSSFSVRKAVQHGKPIFCQASFQ----- 121
 Db 84 IPHSLHSYFIKTSQVDPKPLVIMXIDRIDRGRSFCRVYVYKAVQDGAIFSCQISFHVCMGN 143
 Qy 122 -----QAQSPMQHQFSMTVPVPEELDCELTIDQYLRDNLQ----- 160
 Db 144 GTNKKRTQKPEKPAIRKSSKMEVTPPEQLLPARAALEVLRATKEVDEVTVAGVIOHFL 203
 Qy 161 KRYPLALNRITAAQVEPIEKVNPSPLSOLQRMPEKMPFWVRARYIGGDMKHCVA 220
 Db 204 KEIPDAFERV-----FDRVPNPAYLLKEDTEPMSNTWIRHRENGD-DRLRLHQCVA 256
 Qy 221 YISDYAFAGTALLPHQHKVHFW-VSLDHSWFHA-PFRADHMLYECSPWAGSGRL 278
 Db 257 YLIDLNSMLTAVRPHIRNGCFIPSPSFLSDRCIWNHNEFRIDDMMLYETISSKAGSRAF 316

279 VHGRLWRDQVGLAVTCAQSGVIRKPV 306
 Db 317 IEGRLWRDQVGLAVTCAQSGVIRKPV 343
 RESULT 4
 Q98EG9 PRELIMINARY; PRT: 286 AA.
 AC Q98EG9;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE ACYL-COA THIOESTERASE II.
 GN M14250.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Kishikawa A., Kohikawa A., Takikawa A., Kawashima K., Kimura T.,
 RA Kishikawa Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Motomizu K., Nakayama S., Nakazaki N., Shimizu S., Sugimoto M.,
 RA Tachibana C., Yamada M., Tabata S.,
 RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti".
 RT DNA Res. 7:331-338(2000).
 RL EMBL: AP003003; BAB50950.1; -.
 DR InterPro: IPR003703; Acyl_CoA_thio.
 DR Pfam: PF02551; Acyl_CoA_thio; 1.
 KW Complete proteome.
 SQ SEQUENCE 286 AA; 6177D886273CE322 CRC64;
 Query Match 33.9%; Score 565; DB 16; Length 286;
 Best Local Similarity 41.6%; Pred. No. 1.9e-48;
 Matches 119; Conservative 55; Mismatches 96; Indels 14; Gaps 6;
 Db 22 TVNLNLEPLDELFRGRHYWVPKRLFGQVIGQALVAAKSVSDVHSHCYFVRAGD 81
 Db 10 SLDLERLHNLGRSPQVQWRVFGQVIAQALVAAQRTVPDRVHSLHGYFMRPGD 69
 QY 82 PKLPVLYQVTRTGSFVSRSVRAVQHGKPIFCQASFOQAQSPQHQFSMP-TVPPPP 140
 Db 70 IRVPIVENVDRDGSFTTRVLAIOHQALFESLEASF-OVDERGLEHOFALPDVPPPP 128
 QY 141 EELLDCETLIDQYLDNPKRYPLANRIAQGVPIEIKPNVPSLSQIQRMPEKQMEW 200
 Db 129 EGLPTQQLLER-----AERVPAVERVWARERPLERPVNLQHYESDKLPPRONW 181
 QY 201 VRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPH---OMQHKVHFVWVLDHSMFNAFP 257
 Db 182 IRLAGVDP-DRALGVLLVLSMDTLDTSTFARGCLDFDPDQ-AASLDHSMFVHRHP 239
 QY 258 RADHMLVYECESWAGSGRLVHGRNPDQVGLAVTCAQSGVIRK 303
 Db 240 SLDGWLLYTAQDSSSSGSGRGSGLTARQDGLIASMAQEGILRLK 285
 RESULT 5
 Q91080 PRELIMINARY; PRT: 290 AA.
 AC Q91080;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ACYL COA THIOESTERASE II.
 GN TETR.
 OS Streptomyces coelicolor.
 279 VHGRLWRDQVGLAVTCAQSGVIRKPV 306
 Db 317 IEGRLWRDQVGLAVTCAQSGVIRKPV 343
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=1902;
 [1]
 RP SEQUENCE FROM N.A.
 RA BROWN S.P., Harris D.;
 RA STRAIN=A3(2);
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=A3(2);
 RC Cordero A.M., Parthill J., Barrell J., B.G., Rajandream M.A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RX Redenbach M., Kleser H.M., Denapsaite D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome".
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL163641; CAB87210.1; -.
 DR HSSP: P23911; ICR0.
 DR InterPro: IPR003703; Acyl_CoA_thio.
 DR Pfam: PF02551; Acyl_CoA_thio; 1.
 SQ SEQUENCE 290 AA; 31992 MW; BE7254C37FB82E4E CRC64;
 Query Match 33.7%; Score 562; DB 2; Length 290;
 Best Local Similarity 42.1%; Pred. No. 3.9e-48;
 Matches 123; Conservative 42; Mismatches 113; Indels 14; Gaps 7;
 QY 15 LRSVLVTLMLEPLDELFRGRHYWVPKRLFGQVIGQALVAAKSVSDVHSHCY 74
 Db 5 LOSLL-DLIDLIERIDIEYGRSRSVVPVFGVAAQAAAGRTVPDGRHSHUHA 62
 QY 75 YFVRAGDKPLVLYQVTRTGSFVSRSVRAVQHGKPIFCQASFOQAQSPQHQFSM 134
 Db 63 YFRLPQDGAIVYVDRGRSFTTRVVAVQHGKPIFTLSASQYFEG-IDRQAPM 121
 QY 135 PTVPPEELDCETLIDQYLDNPKRYPLANRIAQGVPIEIKPNVPSLSQIQRM 193
 Db 122 PSAPDPAFLPTGEERLQY---PHLPAE---TVERFLERAAVLDLRYDDPFGDKGTFR 175
 QY 194 EPKQMPVVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPH---OMQHKVHFVWVLDH 250
 Db 176 EPISQVWFRTKGLAD-DPLHVVVLAIVYSDMTLLDSVLLAHGCGHAGVGVGASLDHA 234
 QY 251 WFFHAFRADHMLVYECESWAGSGRLVHGRNPDQVGLAVTCAQSGVIRK 302
 Db 235 WFFHFRFRADHMLVYECESWAGSGRLVHGRNPDQVGLAVTCAQSGVIRK 286
 RESULT 6
 Q9PEK7 PRELIMINARY; PRT: 310 AA.
 AC Q9PEK7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN XP1021.
 GN XP1021.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.C., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Cautão N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferrer J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Krumme J.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Matsukuma A.Y.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Monteiro-Vitorello C.B.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nuan A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto R.B., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Sa J.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.P., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silva A.P., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA Vallada H., Van Sluys M.F., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldin S.J., Zetser J.C.,
 RA "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
 RL Nature 406:151-159(2000).
 DR EMBL: AE003939; AAF83831.1; -;
 DR HSSP: P23911; 1C8U.
 DR InterPro: IPR003703; Acyl_CoA_thio.
 DR Pfam: PF02551; Acyl_CoA_thio; 1.
 DR Hypothetical protein: Complete proteome.
 SK Q9KJ22
 SQ SEQUENCE 310 AA; 35218 MW; GDF0D04B24A4064A CRC64;

Query Match 33.6%; Score 559; DB 16; Length 310;
 Best Local Similarity 39.5%; Pred. No. 8.5e-48;
 Matches 117; Conservative 54; Mismatches 107; Indels 18; Gaps 5;

QY 21 TVLNLEPLEDLEDFRGHYYWPAKRLFGGQIVGOALVAAKSVSDVHSHLCYFVRAG 80
 DB 26 TKMUSLERLEMLFRGNDRDGTGYVFGGQILAAQALAAQNTIENSHLSHLYFLRTG 85
 QY 81 DKLPVLVQVETRTGSSFSVRSKAVQHGKPIFCOASFQOAFSPMOHFSMTVPVPP 140
 DB 86 NIOQPIIYVETRTGSSFSVRSKAVQHGKPIFCOASFQOAFSPMOHFSMTVPVPP 144
 QY 141 EELDCETLIDVLDLPNLOKRYPLALNRIAAQEVPIEKVNPSPLSQLORMEPKQFW 200
 DB 145 EDIAPTPPL-----PPEALPNLQNTWNLGSGFEFRFPNNTLDLQKSHAFOLHM 198
 QY 201 VRANGYIGSDMKHCCVAAIYSDYAFGLTALLPHQKHVHP-----MVSLDRSMWFH 254
 DB 199 LRNAPLGD-RIELHQILTYASDFQLETTATSS-----RGINHTYPOVMASLDHALWFH 253
 QY 255 APPRADHNLVCEESPWAGSGRLVHGRLMRQDGLAVTCAQEGVIRPVQVSGSK 310
 DB 254 RPRFIDWMLLYLESTAGAGLANGQFFRNGVILANTAQEGMLRNTNHTTFSK 309

RESULT 7
 Q9KJ22
 ID Q9KJ22 PRELIMINARY; PRT: 286 AA.
 AC Q9KJ22;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ACYL-COA THIOESTERASE II.
 GN VCI063.
 OS *Vibrio*.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
 OC NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-EL TOP M1961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Haidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Gili S.R., Haft D.H., Hickley E.K., Peterson J.D., Mayday L.A.,
 RA Gil S.R., Nelson K.E., Read T.D., Tattell H., Richardson L.,
 RA Ermolaev M.D., Vatschyan N., Bass S., Olin H., Dragoli I., Sellers P.,
 RA McDonald L.E., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.,
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*,"
 RL Nature 406:477-483(2000).
 RL EMBL: AE004187; AAF94222.1; -;
 DR HSSP: P23911; 1C8U.
 DR TIGR: VC1063; -;
 DR InterPro: IPR003703; Acyl_CoA_thio.
 DR Pfam: PF02551; Acyl_CoA_thio; 1.
 DR Complete proteome.
 SK Q9KJ22
 SQ SEQUENCE 286 AA; 32231 MW; 1FAAF066220D0B00 CRC64;

Query Match 32.6%; Score 543; DB 16; Length 286;
 Best Local Similarity 41.0%; Pred. No. 3.1e-46;
 Matches 118; Conservative 48; Mismatches 100; Indels 22; Gaps 7;
 QY 22 TVLNLEPLEDLEDFRGHYYWPAKRLFGGQIVGOALVAAKSVSDVHSHLCYFVRAG 81
 DB 10 SLIQLEKLEEGLYRGASENLGLPVQVGGVGOALSAAKYTESDRTVHSHSYLYPGD 69
 QY 82 PKLPVLVQVETRTGSSFSVRSKAVQHGKPIFCOASFQOAFSPMOHFSMTVPVPP 141
 DB 70 PEKPIIVDENVLRGKSGFSTRKRAIQNGRPIFYLTATYHGDAPG-FEHQRTMPVVPVGE 128
 QY 142 EELDCETLIDVLDLPNLOKRYPLALNRIAAQEVPIEKVNP-SPLSQLORMEPKQFW 200
 DB 129 NPASESELAQ-----IAHFLPEKLOKAFCEKPIENRVINFLPK-ERAEKPYLYM 181
 QY 201 VRANGYIGSDMKHCCVAAIYSDYAFGLTALLPH-----QKHQKHVHWSLDHSMWF 253
 DB 182 TRANGMDP-DQLIHQYLLGYASDNGELVLTALPHGVSLSMTPKFO-----VATIDHSIWF 235
 QY 254 HAPRADHNLVCEESPWAGSGRLVHGRLMRQDGLAVTCAQEGVIR 301
 DB 236 HRPFKMDMLLPATESATNTRGLVRGEIYDRQNLVATAVQEGVMR 283

RESULT 8
 Q9KJ22
 ID Q9KJ22 PRELIMINARY; PRT: 288 AA.
 AC Q9KJ22;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ACYL-COA THIOESTERASE.
 GN SCG8A.07.
 OS *Streptomyces coelicolor*.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Brown S.P., Harris D.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DDAJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bunley S.D., Parthill J., Barrell B.G., Rajandream M.A.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DDAJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;

```

RA Redenbach M., Kleser H.M., Denepaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb streptococci coelicolic A3(2) chromosome.";
DR Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL353863; CAB88937.1; -.
DR HSSP; P23911; 1C8U.
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
SQ SEQUENCE 288 AA; 32432 MW; 6380F05AAE1DC5B4 CRC64;

Query Match 32.5%; Score 541.5; DB 2; Length 288;
Best Local Similarity 41.3%; Pred. No. 4.4e-46;
Matches 119; Conservative 49; Mismatches 105; Indels 15; Gaps 7;

OY 20 VTTVNLNLEPDLDFGRHYWVPKRLFGQGVQALVAANKSVSDVHVHSLHCYVRA 79
DB 8 LVALLDUEQIEVNIIFGRSPESLQRFVCGVQAGVANGRTDGRPHVSHLHAYRFP 67
OY 80 GPKLPVLVQVETPTGSSFSVRKAVQHGKPTFCOASFQOAPSPMQIFSMPH--V 137
DB 68 GRGPIVTVQVDRDGRSGFTTRVTAQGRITFTASFIKPFQSPENQLP-PRKV 126
OY 138 RPEPELLDCELTLDKLDPNLOKRYPLALNIAAQVPTETKPNPSPLS--OLQRM 195
DB 127 PDPESLTVADVEVREHL-----GALPEQLERMARQ-PEDIRYVDRLSAEDVEGAEP 179
OY 196 KQMFVVRARGYIGEGDMKHCVAAIISDYAFGLTALLPHQ--WQKVFHFMVSLDHSWF 253
DB 180 RSNVMEANGPLGD-DPLVHTCALTYASDWTLLDAVRIPVEPLKPRGPFDMASLDHAWF 238
OY 254 HAPFRADHNLVYECESFWAGSGRLVHGRVQDQGLAVTCAQEGVIR 301
DB 239 HRFRADENFLYDOESPATNGGRLARGRIYDRCRMLVSVVVOEGLFR 286

RESULT 9 PRELIMINARY; PRT; 414 AA.
OY Q95068
AC Q95068;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 46.2 KDA PROTEIN.
GN C37H5.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodidae; Pelodinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology";
RL Science 282:2012-2016(1998).
RB SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Davidson S., Gillam B.;
RT "The sequence of C. elegans cosmid C37H5.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct submision.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88315; AAK68337.1; -.
KW Hypothetical protein.
SQ SEQUENCE 414 AA; 46244 MW; A13CB3AC514A26BF CRC64;

```

```

Query Match 31.1%; Score 518.5; DB 5; Length 414;
Best Local Similarity 37.0%; Pred. No. 1.4e-43;
Matches 121; Conservative 53; Mismatches 116; Indels 37; Gaps 8;

OY 14 DLRSVTVTVNLNLEPDLDFGRHYWVPKRL---ECGTVGQALVAANKSVSDVHVH 70
DB 95 DTRAGLIDTFLNLRVDKNLILARLLKGNLSLPVYGGQVIGQALSATATVEGVFVN 154
OY 71 SLIKYFVRADGPKFLVLYQVETPTGSSFSVRKAVQHGKPTFCOASFQOAPSPMQH 130
DB 155 SLHSYFQSGNVERPILYQVDRIRKGSFCTRLVKALDCEALFTVQLSFHREADSI 214
OY 131 QESNPTVPPEELDCETLIDQYLRDNL-----OKRYPLALNIAAQVPEI 179
DB 215 OLPAPEVPAPDSLESDLTIFERIKKANIPPAALAMIGFKQKEIPAFERI-----FSF 268
OY 180 KPVNPSPLSQLORMP-----KQMFVVRARGYIGEGDMKHCVAAIISDYAF 228
DB 269 RPVDIDSYLCKDKDHTAGHGHTDAYRSVYLIKANENIGD-DPLRLAAATISDATMI 327
OY 229 GTALLPHQHKVHM-VSLDHSWFHAP--FRADHNLVYECESFWAGSGRLVHGRV 286
DB 328 ETALRPHSKRGFIPSMALTLDHSIMNHTDNFRVDWMLYENHSTIAGGRSLIEGKLW 387
OY 287 DGLAVTCAQEGVIRV---KPOVSEK 310
DB 388 DGRLVFTQELINTAKNRNPASPAK 414

RESULT 10 PRELIMINARY; PRT; 284 AA.
OY Q9A2B6
AC Q9A2B6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ACYL-COA THIOESTERASE II.
GN CC3649.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089; CB15;
RX MEDLINE=21173681; PubMed=11259647;
RA Nierman W.C., Feildberg J.T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Hildebrand J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Decker T., Nelson R.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Kolonay J.F., Smit J.J., Durkin A.S., Winn M.L., Haft D.H.,
RA Klotzel J., Tran K., Wolf A., Vamathevan J., Fraser C.M.;
RA Salzbarg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE006023; AAK25611.1; -.
DR HSSP; P23911; 1C8U.
DR TIGR; CC3649; -.
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
KW Complete proteome.
SQ SEQUENCE 284 AA; 31518 MW; F65A4C87C5560AF7 CRC64;

Query Match 30.1%; Score 501.5; DB 16; Length 284;
Best Local Similarity 41.5%; Pred. No. 4.4e-42;
Matches 118; Conservative 54; Mismatches 99; Indels 13; Gaps 9;

OY 23 VLNLEPDLDFGRHYWVPKRLFGQGVQALVAANKSVSDVHVHSLHCYVRA 82
DB 10 ILDLPEIFVNLFRGVSPNDGPFRIFGVLVTAQALLAARTVPDRV-CHSLHAYTRPGDV 68

```

QY 83 KLPVLYOVERTRTGTSFVSRSKAYQHKPFIICQASFOQAQSPMOHOFSMPTVPPPEE 142
 DB 69 TAPVLYEVARADGGTFTTRVAATQHGQIFNLAASFQTPEDG-FEHOSEMPASVDPE 127
 QY 143 LDCETLIDQYLRDNPLOKRYPLALNRIAAQEVPIETKPNP-SPLSOLQRMPEKMPFW 201
 DB 128 L---PTAD-FLR--SLGDDIHPKMAIARRPVDIRWDIPQNDIAPVKSKTQV-WM 180
 QY 202 RARGVIGEDMKHCCVAAVYISDYAFGLTALLPHQ--WQHKVHEMVSLDHSMFHPAPRA 259
 DB 181 RAKAPLGD-DVMQQAALAYASDMAFMSALRPHGLIWTTPGTOAASLDHAMFWHPHPNF 239
 QY 260 DHWMLYCESPWAGSGRLVHGLRLRQDQGLAVTCAQEGVIRVK 303
 DB 240 NDWTLPADQSPASQSGRLVQGMFSQDGRLLASVAQGLMVR 283

RESULT 11
 ID Q9C321 PRELIMINARY; PRT: 320 AA.
 AC Q9C321
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE TIN6.10 PROTEIN.
 GN TIN6.10
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA.
 RA Liu S.-C., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,
 RA Toriumi M., Chin C., Chlou J., Choi E., Chung M., Gonzalez A.,
 RA Hwang B., Liu A., Vaysberg M., Altfai H., Brooks S., Buehler E.,
 RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,
 RA Kim C., Lam B., Miranda M., Nguyen N., Palm C., Shinn P.,
 RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;
 RA Submitted (JUN-2000) to the EMBL/GenBank/DBS databases.
 RL EMBL; AC009273; AAF78401.1; *.
 DR HSP; P23911; IG8U.
 DR InterPro; IPR003703; Acyl_CoA_thio.
 DR Pfam; PF02551; Acyl_CoA_thio; 1.
 NCBI_SEQUENCE 320 AA; 36157 MW; 03A82A9G33E1BFC CRC64;

Query Match 30.0%; Score 499.5; DB 10; Length 320;
 Best Local Similarity 37.1%; Pred. No. 8.2e-42;
 Matches 115; Conservative 49; Mismatches 101; Indels 45; Gaps 7;

QY 33 LFRGRHVVY-PAK-----RIFGQIVGQALVAARKSVSDVHVSUHCY 75
 DB 2 LPFGTHATFRFRTGRTIPLDAPIFGQVFGQFVQALAAAKTVDFLKVHSLHST 61
 QY 76 FVRAGD-----PKLPVLYOVERTRTGTSFVSRSKAYQHKPFIICQASFOQAQSP 127
 DB 62 FLVAGDIDSLLLAVPTIYQVHRIDGNFATRVDAVQGNILFILLASFQEQGF 121
 QY 128 MOHGSMTVPPPELDCETLIDQYLRDNPLOKRYPLALNRIAAQEV--PIETKPNP 184
 DB 122 EHOESTMFSVPDPTLLSLEELRESRTIDPHLPSY---RNKVAIRNFYVPIIRFCBP 178
 QY 185 SPLSOLQRMPEKMPFWARVYIGEDMKHCCVAAVYISDYAFGLTALLPHQHKVHEM 244
 DB 179 SNTNQTSPRLNRYWFRAKGRISD-DQALHRCVAFASDLIFCGVGLNPHRRKYSAA 237
 QY 245 VSLDHSMMFHPAPRADHMLY-----ECESPWAGSGRLVHGLRLRQDQGLA 291
 DB 238 LSLDHAMFWHRLRADENLLYVRFTHLQHGKCCIQTIVSPTAHETRGVTFGQVNRKGLV 297

QY 292 VTCAQEGVIR 301
 DB 298 VSLTQEAALLR 307

RESULT 12
 ID Q9C321 PRELIMINARY; PRT: 292 AA.
 AC Q9C321
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE TESB.
 GN TESB OR PM0570.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OC NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May H.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kepur V.;
 RA Complete genomic sequence of Pasteurella multocida Pm70.*;
 RL J. Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
 DR EMBL; AC006694; AAK02654.1; *.
 DR HSP; P23911; IG8U.
 DR Pfam; PF02551; Acyl_CoA_thio.
 DR Complete Proteome.
 NCBI_SEQUENCE 292 AA; 32977 MW; CB86741442225AF1 CRC64;

Query Match 27.8%; Score 463; DB 16; Length 292;
 Best Local Similarity 38.8%; Pred. No. 3.4e-38;
 Matches 114; Conservative 48; Mismatches 114; Indels 18; Gaps 8;

QY 23 VINLEPLEDELFGRHVVYWPARKLFGQIVGQALVAARKSVSDVHVSUHCYFVRAGDP 82
 DB 11 LQLERLDDFLFRGKSDGLGRVFGGVQAQALSNATQVAPDRVLHSCHAYFLSPGDS 70
 QY 83 KLPVLYOVERTRTGTSFVSRSKAYQHKPFIICQASFOQAQSPMOHOFSMPTVPPPEE 142
 DB 71 QHGIIYOVETLREGNFTALRVKALQHNPEICHITASFQAKBQG-FDHQSKMPDESQ 129
 QY 143 LDCETLIDQYLRDNPLOKRYPLALNRIAAQEVPIETKPV---NPSPLSOLQRMPEKMP 199
 DB 130 LI-AESATIQ-----KLAPHIPEAIRAKFTAEPRPDTRTKYNNPFHGSQ-----PAEQ 179
 QY 200 -WVARVYIGEDMKHCCVAAVYISDYAFGLTALLPHQ--WQHKVHEMVSLDHSMFHPAP 256
 DB 180 AMVKANGHAPO-DHKIQOCLLAYFSDPHCLLTALHPHQKGFQPGMKVATIDHSIWFHRP 238
 QY 257 FRADHMLYCESPWAGSGRLVHGLRLRQDQGLAVTCAQEGVIRVKPQVSESK 310
 DB 239 FDLNDMLLYATESNAPFAARGLARGQIFDROGRLTATTOQEGRLYTPKESEK 292

RESULT 13
 ID Q9C321 PRELIMINARY; PRT: 368 AA.
 AC Q9C321
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE ACYL-COA THIOESTERASE.
 GN TESB
 OS Chlolobolus heterotrophus (Drechslera maydis).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; Cochliobolus.
 OC NCBI_TaxID=5016;
 RN [1]

RP SOURCE FROM N.A.
RA "u.S.W., Turgeon B.G., Yoder O.C.;
RT "Cochitilobius heterostrophus acyl-CoA thioesterase gene TES1.";
DE EMBL: AF323079; GenBank: U000760; the EMBL/GenBank/DBJ databases.
DR HSP: F233187; ICBU:
DR InterP: IPR003703; Acyl_CoA_thio.
DR Pfam: PF02551; Acyl_CoA_thio.1
SQ SEQUENCE 368 AA; 41080 MW; 601D9B787C23BD5A CRC64;

Query Match 27.0%; Score 449.5; DB 3; Length 368;
Best Local Similarity 32.6%; Pred. No. 1e-36;
Matches 117; Conservative 44; Mismatches 107; Indels 91; Gaps 11;

QY 11 PGCDLRSVLTVTNLNLEPL---DEDLERG-RHYWVP--AKRLFGQIVGQALVAAKASVS 64
DB 8 PRADHSQSAENLVLTQLADIDPNIPTNPLHPHGARGINGGAALQTLNNAQTVD 67
QY 65 EDVHVSLHCYFVRAGDPKPLVLYOVERTGSSFSVRSKAVQHKPFIICQASF--Q 121
D 68 PDFTVSHMHCYILAGNSEPTVYHVERVRSKGFATRTVOARGNVIFTTTMSFVRQN 127
QY 122 QAQSPMQHOFNPTVPPEELDCETLIDYLRDPNLQKRYPLALNRIAAQVPIEIKP 181
DB 128 SGGAQVHEHYMPDPVPAPEGSD-----DLKTPNDGQ-----SPFQTLPLIENAD 174
QY 182 VNPSPLSQLQRMKQMFWRVARGYIG-EGDKMHCCVAAIYSDYAFLGTA-----LLPH 235
DB 175 DSDKPHTKKQ-----WIKARGKISPGAGHEAHLISAIAIYMSDSYFICTVARAHKLLRY 228
QY 236 QMQHK----- 240
DB 229 SNQKRSRRSISDEDVLKLEMDAELQRFSEVNSDKQIRLKAELAKSGDAKPE 288
QY 241 VHEVYSLDHSMMFHAP--FRADHMLYECESPMAGSGRLVGHRLWRQDQGLVATCAQE 297
DB 289 IGMWVSLDHTIYFNKPSFRADHEFTMETPWAQDGRGLVSQRMYTKDSLTIASCVOE 347

RESULT 14
O06209 PRELIMINARY; PRT; 281 AA.
AC O06209
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TESB.
GN TESB2 OR RV2605C OR MTCY01A10.28.
OC Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID:1773;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN-H37RV; PubMed:9634230;
RX MEDLINE-96295987; PubMed-9634230;
RA Cole S.F., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigameier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Secherre K., Skellern S., Squares R.,
RA Sulston J.E., Taylor A., Whitthead S., Barrrell B.G.;
RT "Defining the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL NATURE 393:537-544 (1998).
DR EMBL: Z95387; CAS08615.1;
DR HSP: F233187; ICBU:
DR TubercuList; RV2605C;
DR InterP: IPR003703; Acyl_CoA_thio.
DR Pfam: PF02551; Acyl_CoA_thio.1.

KW Complete proteome.
SQ SEQUENCE 281 AA; 31569 MW; D78459281EBC8949 CRC64;

Query Match 26.8%; Score 442.5; DB 16; Length 281;
Best Local Similarity 36.8%; Pred. No. 3.6e-36;
Matches 105; Conservative 46; Mismatches 113; Indels 23; Gaps 7;

QY 20 VTYVNLNLEFLDEDLERGHYHWPFA--KRLFGQIVGQALVAAKASVSDVHVSILKCYF 76
DB 3 IBEIIDLLEQLENNYIRGSVFSPEGFLQRTFGGHVAGSLYSNVRIVDPNVAHSLGYP 62
QY 77 VRAGDPKPLVLYOVERTGSSFSVRSKAVQHKPFIICQASFQQAQSPMQHOFNPT 136
DB 63 LRPGDAKERTVLYERIDGSSCTRRYNAVQHGFIISMAASFQTEGEG-IHODYMPA 121
QY 137 VPPPEELDCETLIDYLRDPNLQK--RYPALNRIAAQVPIEIKPNPSPILSQORME 194
DB 122 APPDGLPLGNSI--KVPDAGFPROFDEWDVCI-----VPRERLLLPCKASQ----- 167
QY 195 PKOMFWVRARGYIGSGDKMHCCVAAIYSDYAFGLTALLPHQHKVHPMWVSLDHSMMFH 254
DB 168 --QVWLHRHRDPLP-DPVLICALAYMSDLTLGSAQVNHLDVDRDQLQVASLDHAMWFM 224
QY 255 APPRADHMLYECESPMAGSGRLVGHRLWRQDQGLVATCAQEGVIR 301
DB 225 RPFRADEMLLYDQSSPSAGSGRALTRGEITFRSGENVAAVNQELTR 271

RESULT 15
O091A5 PRELIMINARY; PRT; 346 AA.
AC O091A5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOPHETICAL 39.7 KDA PROTEIN.
GN C17C3.1.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID:6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None.
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
KA Du Z.;
RT "The sequence of C. elegans cosmid C17C3.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U41279; AAK31429.1;
DR HSP: F233187; ICBU:
DR InterP: IPR003703; Acyl_CoA_thio.
DR Pfam: PF02551; Acyl_CoA_thio.1.
OX Hypochemical protein.
KW SEQUENCE 346 AA; 39688 MW; CF853F9CAF093B2 CRC64;

Query Match 26.4%; Score 440.5; DB 5; Length 346;
Best Local Similarity 35.2%; Pred. No. 7.5e-36;
Matches 106; Conservative 40; Mismatches 94; Indels 61; Gaps 9;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:18:29 ; Search time 30.4 Seconds
(without alignments)
1136.314 Million cell updates/sec

Title: US-09-766-366-1

Perfect score: 1666

Sequence: 1 MGRVATTAALPGDLRSVLV.....VTCAQGVIRKPVQSESKL 311

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
A_Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1666	100.0	311 19	AAW58530 Human peroxisomal thioesterase
2	1666	100.0	311 20	AAW08004 Human PXTE protein
3	1666	100.0	311 22	AAW77935 Protein sequence o
4	1666	100.0	311 22	AAW73420 Human peroxisomal
5	1284	77.1	331 22	ABG20246 Novel human diagno
6	628.5	37.7	270 22	AAU04004 Thioesterase type
7	553	33.2	286 20	AAV08619 E. coli Tili prote
8	489	29.4	271 21	AAW24548 Arabidopsis thalia
9	489	29.4	377 21	AAW24547 Arabidopsis thalia
10	489	29.4	365 21	AAW24546 Arabidopsis thalia
11	484.5	29.1	106 22	ABG20246 Novel human diagno

12	466.5	28.0	282	22	AAG91582	C glutamicum prote
13	466.5	28.0	282	22	AAW76590	Corynebacterium g1
14	466.5	28.0	282	22	AAW76591	Corynebacterium g1
15	449.5	27.0	368	22	AAU03392	Cochliobolus heter
16	433	26.0	298	22	AAU62058	Propionibacterium
17	426	25.6	104	22	ABG20245	Novel human diagno
18	273	16.4	349	20	AAV08620	Yeast CoA thioeste
19	201	12.1	81	22	ABG20244	Novel human diagno
20	196	11.8	140	22	AAU65800	Propionibacterium
21	180.5	10.8	420	22	ABG25236	Novel human diagno
22	145	8.7	91	22	AAW80886	Lipid degradation
23	145	8.7	91	22	AAW83336	P patens lipid met
24	127	7.6	43	22	ABG20247	Novel human diagno
25	111.5	6.7	557	22	ABG26202	Novel human diagno
26	109.5	6.6	591	22	AAW39282	Human polyptide
27	109.5	6.6	594	22	AAW1068	Human polyptide
28	100	6.0	180	18	AAW23719	Platenolide syntha
29	100	6.0	1580	18	AAW22609	Drosophila melanog
30	91	5.5	1425	22	ABW5506	Drosophila sp. mus
31	91	5.5	1425	22	AAE04383	Drosophila sp. mus
32	90.5	5.4	1234	22	AAE07171	Human HuIFRG-3 pro
33	90.5	5.4	1268	22	ABW11796	Human secreted pro
34	89	5.3	2240	15	AAW57819	Acetyl CoA carboxy
35	89	5.3	2325	17	AAW05590	Maize acetyl CoA c
36	89	5.3	2325	19	AAW56736	Maize ACCase enzym
37	89	5.3	2325	21	AAW44667	Maize acetyl CoA c
38	86.5	5.3	1743	22	AAU03388	Cochliobolus heter
39	86	5.2	1718	22	AAU28246	Novel human secret
40	86	5.2	1816	22	AAW80877	Drosophila melanog
41	85.5	5.1	1070	18	AAW08747	Human colon carcin
42	85	5.1	361	22	ABW70969	Drosophila melanog
43	85	5.1	746	22	ABW69572	Drosophila melanog
44	84	5.0	648	22	AAU02193	Transcriptional fac
45	83.5	5.0	331	22	AAU35206	Enterococcus faeca

ALIGNMENTS

RESULT 1

AAW58530
ID AAW58530 standard; Protein: 311 AA.

XX AC AAW58530;

XX DT 01-SEP-1998 (first entry)

XX DE Human peroxisomal thioesterase.

XX DE Human peroxisomal thioesterase.

XX KW Human; peroxisomal thioesterase; PXTE; diagnosis: inflammatory disease;

XX KW fatty acid metabolism disorder; cancer; drug screening.

XX OS Homo sapiens.

XX PN US5776753-A.

XX PD 07-JUL-1998.

XX PF 11-JUN-1997; 97US-0872784.

XX PR 11-JUN-1997; 97US-0872784.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Corley NC, Hillman JL, Shah P;

XX DR WPI; 1998-398039/34.

XX DR N-PSDB; AAV31109.

XX PT New nucleic acid encoding human peroxisomal thioesterase - useful

XX PT for diagnosis, treatment and prevention of fatty acid metabolism

XX PT disorders, cancer and inflammatory disease, and in drug screening

XX

PS Claim 9; Fig 1; 30pp; English.

XX The present sequence represents human peroxisomal thioesterase (PXTE).
 CC The cDNA encoding PXTE was isolated from a human BRAINOT09 clone 2150905
 CC cDNA library. PXTE and its agonists are used in treatment, diagnosis
 CC and prevention of disorders of fatty acid metabolism e.g. neuronal
 CC disorders, Zellweger or Down's syndrome, Refsum's disease, epilepsy,
 CC Alzheimer's disease, depression, tardive dyskinesia, multiple sclerosis,
 CC Parkinson's disease, schizophrenia and Tourette's syndrome. Fragments of
 CC PXTE DNA can also be used for diagnosis and monitoring (in hybridisation
 CC or amplification methods), including determining a predisposition to
 CC disease, for gene mapping and to detect PXTE-encoding nucleic acid,
 CC particularly after amplification. PXTE can be used to raise antibodies
 CC (Ab) used therapeutically, in competitive drug screens, as immunoassay
 CC reagents for diagnosis and monitoring of diseases and for purification of
 CC natural PXTE, and to screen for specific binding agents antagonists e.g.
 CC specific Ab, antisense sequences or ribozymes which are used in treatment
 CC of cancers and inflammatory diseases (e.g. adult respiratory distress
 CC syndrome, asthma, diabetes mellitus, microbial and other infections).
 CC Administration of proteins, agonist or antagonists, optionally combined
 CC with conventional pharmaceuticals providing a synergistic effect, is
 CC e.g. orally, by injection or topically.

XX Sequence 311 AA;

Query Match 100.0%; Score 1666; DB 19; Length 311;
 Best Local Similarity 100.0%; Pred. No. 4.5e-177;
 Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGRAVATAALPPGDLRSVLVTVNLPLEDLEDFRGRHYWVPAKRLFGGQIVGQALVAAA 60
 Db 1 mgravataaalppgdrlsvltvtvnlpledledfrgrhywvpakrlfggqlvgqalvaaa 60
 Qy 61 KVSSEDVHVHSLHCYFVRAGDPKLPVLYQVTRTGSFVSRSVAKVHGKPIFTCAQSF 120
 Db 61 kvsedvvhvshlhcylvragdplkplvyqverttgsfvsrsvskavghkpfiftcqsaf 120
 Qy 121 QQAQPSPMQHOFSMPTVPPPELDCETLIDQYLRDNLQKRYPLALNRIAAQGVPIEIK 180
 Db 121 qqaqpspmqhofsmptvpppeelldcetliidqyldrdnlqkryplalnriaaevpieik 180
 Qy 181 PVNPSPLSQLRMEPKQMFVWRARGYIGEDMKHCCVAAVYISDYAFGLTALLPHQWQHK 240
 Db 181 pvnpsplsqlrmeqkmtfwrargyigedmknhccvvaayisdyaflgtallphqwgkh 240
 Qy 241 VHFVWSLDHSMWFIAPRADHMLYECESFWAGGSRGVLVHGRMQDGLVAVTCAQEGVI 300
 Db 241 vhfvwslhsmwfiapradhmlyecesfwaggsrglvhgrmqdglvavtcaqegvi 300
 Qy 301 RVKPVQVSESKL 311
 Db 301 rvkpqvseskl 311

RESULT 2
 ID AAY08604 standard; Protein; 311 AA.

XX AAY08604;

XX 05-AUG-1999 (first entry)

XX Human PXTE protein.

XX PXTE; peroxisomal thioesterase; treatment; disorder; biosynthesis;
 KW fatty acid metabolism; chain termination; fatty acid; coenzyme deficiency;
 KW inflammation; adrenoleukodystrophy; Refsum's disease; mental retardation;
 KW disordered neuronal migration; hypotonia; tapetoretinal degeneration;
 KW sensorineural hearing loss; cystic change; kidney; skeletal change.
 OS Homo sapiens.

XX

PN US5911984-A.
 XX
 PD 15-JUN-1999.
 XX
 PP 19-JUN-1998; 98US-0100851.
 XX
 PR 11-JUN-1997; 97US-0872784.
 XX
 PR 19-JUN-1998; 98US-0100851.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Corley NC, Hillman JL, Shah P;
 XX WPI; 1999-357169/30.
 XX
 DR N-PSDB; AAX77452.
 XX

Peroxisomal thioesterase enzymes useful for treating disorders
 associated with fatty acid metabolism

Claim 1; Fig 1A-C; 29pp; English.

XX This invention describes a novel human peroxisomal thioesterase enzyme
 CC (PXTE) and its use to treat disorders associated with fatty acid
 CC metabolism. Thioesterases catalyse the final and rate-limiting step in the de
 CC novo biosynthesis of fatty acids. PXTE may be used to treat cancer,
 CC inflammation, disorders associated with fatty acid metabolism, and the
 CC expression of PXTE such as adrenoleukodystrophy Refsum's disease, and
 CC PXTE deficiency. Patients suffering from defective peroxisomal fatty
 CC acid metabolism exhibit disordered neuronal migration, hypotonia, mental
 CC retardation, tapetoretinal degeneration, sensorineural hearing loss,
 CC cystic changes in the kidney, skeletal changes and death.

XX Sequence 311 AA;

Query Match 100.0%; Score 1666; DB 20; Length 311;
 Best Local Similarity 100.0%; Pred. No. 4.5e-177;
 Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGRAVATAALPPGDLRSVLVTVNLPLEDLEDFRGRHYWVPAKRLFGGQIVGQALVAAA 60
 Db 1 mgravataaalppgdrlsvltvtvnlpledledfrgrhywvpakrlfggqlvgqalvaaa 60
 Qy 61 KVSSEDVHVHSLHCYFVRAGDPKLPVLYQVTRTGSFVSRSVAKVHGKPIFTCAQSF 120
 Db 61 kvsedvvhvshlhcylvragdplkplvyqverttgsfvsrsvskavghkpfiftcqsaf 120
 Qy 121 QQAQPSPMQHOFSMPTVPPPELDCETLIDQYLRDNLQKRYPLALNRIAAQGVPIEIK 180
 Db 121 qqaqpspmqhofsmptvpppeelldcetliidqyldrdnlqkryplalnriaaevpieik 180
 Qy 181 PVNPSPLSQLRMEPKQMFVWRARGYIGEDMKHCCVAAVYISDYAFGLTALLPHQWQHK 240
 Db 181 pvnpsplsqlrmeqkmtfwrargyigedmknhccvvaayisdyaflgtallphqwgkh 240
 Qy 241 VHFVWSLDHSMWFIAPRADHMLYECESFWAGGSRGVLVHGRMQDGLVAVTCAQEGVI 300
 Db 241 vhfvwslhsmwfiapradhmlyecesfwaggsrglvhgrmqdglvavtcaqegvi 300
 Qy 301 RVKPVQVSESKL 311
 Db 301 rvkpqvseskl 311

RESULT 3

ID AAG77835 standard; protein; 311 AA.

XX AAG77835;

XX 03-JAN-2002 (first entry)

XX Protein sequence of human peroxisomal thioesterase (PXTE).

XX KW Human peroxisomal thioesterase; PXT; immunogenic; cytostatic;
 KW neotropic; neuroleptic; antiinflammatory; neuroprotective;
 KW cerebroprotective; Down's syndrome; gene therapy; cancer; inflammation;
 KW fatty acid metabolism disorders; neuronal disorders; Alzheimer's disease;
 KW amnesia; Huntington's disease; cerebropatarenal syndrome;
 KW hybridisation probe; adrenoleukodystrophy; schizophrenia;
 KW adrenoleukodystrophy; Crohn's disease; adenocarcinoma; leukaemia;
 KW Addison's disease; adult respiratory distress syndrome.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 XX Region 309..311
 XX /note= "C-terminal SKL (ser-lys-leu) matrix targeting
 XX sequence"
 XX US01024786-A1.
 XX 27-SEP-2001.
 XX 18-JAN-2001; 2001US-0766366.
 XX 11-JUN-1997; 97US-0872784.
 XX 19-JUN-1998; 98US-0100851.
 XX 09-MAR-1999; 99US-0265294.
 XX (INCY-) INCYTE PHARM INC.
 XX Hillman JL, Shah P, Corley NC;
 XX WPI; 2001-638498/73.
 XX N-PSDB; AAF78821.
 XX New human peroxisomal thioesterase polypeptide and polynucleotide,
 XX useful in diagnosing, preventing and treating cancer, inflammation and
 XX disorders associated with fatty acid metabolism -
 XX Claim 1; Fig 1; 33pp; English.
 XX The protein sequence represents the human peroxisomal thioesterase
 XX enzyme (PXT) of the invention. The invention also includes the
 XX polynucleotide sequence coding for this protein. Also the invention
 XX provides expression vectors, host cells, agonists, antibodies and
 XX antagonists. The invention also provides methods for treating disorders
 XX associated with expression of PXT. The invention has the following
 XX activities of being cytostatic, antiinflammatory, neuroprotective,
 XX neotropic, cerebroprotective and neuroleptic although no supporting data
 XX is given. The PXT polypeptides are useful in diagnosing, preventing and
 XX treating cancer (e.g. adenocarcinoma or leukaemia), inflammation (e.g.
 XX colon and small intestine afflicted by Crohn's disease) and disorders
 XX associated with fatty acid metabolism, such as neuronal disorders
 XX including Alzheimer's disease, amnesia, Down syndrome, Huntington's
 XX disease, or schizophrenia by using the PXT polypeptides in gene therapy.
 XX In one embodiment, an antagonist of PXT may be administered to a subject
 XX to prevent or treat inflammation caused by conditions such as Addison's
 XX disease or adult respiratory distress syndrome. The PXT polynucleotides
 XX can be used to detect and quantify gene expression in biopsied tissue. In
 XX the diagnosis of conditions, disorders or diseases associated with PXT
 XX expression e.g. adrenoleukodystrophy, adrenomyeloneuropathy, or
 XX cerebropatarenal syndrome, to detect activation or induction of various
 XX cancers, to generate hybridization probes useful for mapping naturally
 XX occurring genomic sequences, and for screening libraries in various drug
 XX screening techniques.
 XX Sequence 311 AA;

Query Match 100.0%; Score 1666; DB 22; Length 311;
 Best Local Similarity 100.0%; Pred. No. 4.5e-177;
 Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRVATAALPPGDLRSVLVTTVINLEPDLDFRGRHWVPAPKLFGGQIVGQALVAAA 60

Db 1 mdravataaalppgdlrsvlvttvinlepldedlfrgrhvwvpakrlfvgqivgqalvaa 60
 QY 61 KSVSEDEVHSHLCHYFVRAGDPKLPVLYQVETRTGSSFSVRSKAVGRKPIFICQASF 120
 Db 61 ksvsedvvhshlchfyvradgpklpvlyqvertrtgssfsrvskavdghpfcicqasf 120
 QY 121 QQAQSPMQHQSMTPTVPPEELDCETLIDQYLDPNLQKRYPLALNRIAAQEVPIEK 180
 Db 121 qqaqspmqhqsmtptvppeelldcetldqylrdpolkryplalnriaaevpiek 180
 QY 181 PYNPSPISQIQRMEPKQNFVVRARYIGESDMKHCVCVAAVISDYAFIGTALLPHOMOHK 240
 Db 181 pynpbspisqlqrmpkqnfvrarvigedmkhmcvvaayisdyafigtallphqgdhk 240
 QY 241 VHPWVSLDHSMWFHAPPRADHMLYECESPWAGSGSLVHGRIWRQDGYLVAYCAQEGVT 300
 Db 241 vhtmwslshsmwfhapiradhmlylceespwagsgslvghrwrqdgylvaycaqdegvt 300
 QY 301 RVKPOVSESKL 311
 Db 301 rvkpvaseeskl 311
 RESULT 4
 AAB73420
 ID AAB73420 standard; Protein; 311 AA.
 AC AAB73420;
 DT 02-JUL-2001 (first entry)
 DE Human peroxisomal thioesterase, PXT.
 KW Human; peroxisomal thioesterase; PXT; fatty acid biosynthesis;
 KW chain termination; detection; diagnosis; monitoring; hybridisation;
 KW inflammation; cancer; defective fatty acid metabolism; neuronal disorder;
 KW adrenoleukodystrophy; adrenomyeloneuropathy; cerebropatarenal syndrome;
 KW Zellweger syndrome; Refsum's disease; Alzheimer's disease; amnesia;
 KW amyotrophic lateral sclerosis; bipolar disorder; Huntington's disease;
 KW Down's syndrome; multiple sclerosis; neurofibromatosis; schizophrenia;
 KW Parkinson's disease; paranoid psychosis; Tourette's syndrome;
 KW drug screening.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 XX Region 309..311
 XX /note= "peroxisomal matrix targeting sequence"
 XX US6210890-B1.
 XX 03-APR-2001.
 XX 09-MAR-1999; 99US-0265294.
 XX 11-JUN-1997; 97US-0872784.
 XX 19-JUN-1998; 98US-0100851.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Hillman JL, Shah P, Corley NC;
 XX WPI; 2001-289629/30.
 XX N-PSDB; AAF76400.
 XX Detecting human peroxisomal thioesterase (PXT) in a biological sample,
 XX useful for diagnosing disorders associated with fatty acid metabolism,
 XX comprises hybridizing a PXT polynucleotide complement with the sample
 XX nucleic acid -
 XX Claim 1; Fig 1A-C; 30pp; English.

CC This sequence represents a human peroxisomal thioesterase, pXTE.
 CC pXTE, like other thioesterases, catalyzes the chain-terminating
 CC step in the de novo biosynthesis of fatty acids. pXTE is expressed
 CC in neuronal, gastrointestinal and secretory tissues, and cells associated
 CC with inflammation and cancer. The invention relates to a method of
 CC detecting nucleic acids encoding pXTE involving the hybridisation of a
 CC pXTE nucleic acid probe to pXTE nucleic acids in a sample, and the
 CC subsequent detection of the hybridisation complex. The method is useful
 CC for the diagnosis of cancer, inflammation and disorders associated with
 CC fatty acid metabolism, and may also be used to monitor the treatment of
 CC a patient with a disorder associated with pXTE expression. Disorders
 CC associated with defective fatty acid metabolism include neuronal
 CC disorders such as adrenoleukodystrophy, adrenomyeloneuropathy,
 CC cerebrohepato renal syndrome (Zellweger syndrome), Refsum's disease,
 CC Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar
 CC disorder, Huntington's disease, Down's syndrome, multiple sclerosis,
 CC neurofibromatosis, Parkinson's disease, paranoid psychoses,
 CC schizophrenia, or Tourette's syndrome. In addition to their use as
 CC probes in the method of the invention, pXTE nucleic acids may also
 CC be used in chromosomal mapping, detecting polymorphisms associated
 CC with disease, in gene therapy, and in drug screening techniques. The
 CC pXTE protein may also be used in drug screening, and to raise
 CC antibodies.

XX Sequence 311 AA;

Query Match 100.0%; Score 1666; DB 22; Length 311;
 Best Local Similarity 100.0%; Pred. No. 4.5e-177;
 Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGRAVATAALPGDLSRLSVTVTVNLNLEPLDEDLFRGRHYWVPKRLFGGQIVGQALVAAA 60

Db 1 mgravataalpddlsrlsvltvtvnllepldedlfrgrhywvpakrlfsggqivgqalvaa 60

Qy 61 KSVSEDDVHVSILCHYFVRAGDPKLPVLVQVVERTTSSFSVRSVKAVOHGKPIFTCOASF 120

Db 61 kaveddvhvsilhcyfvrsgdpkplvlyqvverttsgsfsvrskvavghgkpficqasf 120

Qy 121 QQAQPPSMQHOFMSMPTVPPPELLDCEITLIDQYLDPNLQKRYPLALNRIAAQEVPIETK 180

Db 121 qqaqpgpmghfsmptvpppeellidcetlidyldpnlqkryplalnriaaqevpieik 180

Qy 181 PVPNPSLQOLRMEPKMFWRNARGYIGEGDNKHCVCVAYISDYAFELGTALLPHQOHK 240

Db 181 pvpnpplslqrlmepkqmfwrargyigegdmkhcvcvayisdyafigtaliphqwhk 240

Qy 241 VHPMVSLSHWNHFAPRADHMLIYECESPMWAGSGRLVHGLWRQDGVILVATCAQEGVI 300

Db 241 vhmfvslshwnhfapradhmllyecespmwagsrglvhglwrqdgvlvlatcaqegvi 300

Qy 301 RVKQVQSVESKL 311

Db 301 rvkqvqvseskl 311

RESULT 5

ABG20249

Id ABG20249 standard; Protein: 331 AA.

XX AC ABG20249;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #20240.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO2001/75067-A2.

XX

PD 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX MPI; 2001-639362/73.
 XX N-PSDB; AAS84436.

XX New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity.

XX Claim 20; SEQ ID NO 50608; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 XX polypeptide (II) sequences (I) is useful as hybridisation probes,
 XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 XX and gene mapping, and in recombinant production of (II). The
 XX polynucleotides are also used in diagnostics as expressed sequence tags
 XX for identifying expressed genes. (I) is useful in gene therapy techniques
 XX to restore normal activity of (II) or to treat disease states involving
 XX (II). (II) is useful for generating antibodies against it, detecting or
 XX quantitating a polypeptide in tissue, as molecular weight markers and as
 XX a food supplement. (II) and its binding partners are useful in medical
 XX imaging of sites expressing (II). (I) and (II) are useful for treating
 XX disorders involving aberrant protein expression or biological activity.
 XX The polypeptide and polynucleotide sequences have applications in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits to assess biodiversity
 XX and to produce other types of data and products dependent on DNA and
 XX amino acid sequences. ABG00010-ABG30377 represent novel human
 XX diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 331 AA;

Query Match 77.1%; Score 1284; DB 22; Length 331;

Best Local Similarity 81.7%; Pred. No. 2e-134; Indels 12; Gaps 12;

Matches 263; Conservative 11; Mismatches 36;

Qy 2 GRAVATAALPGDLSRLSVTVTVNLNLEPLDEDLFRGRHYWVPKRLFGGQIVGQALVAAA 61

Db 10 ggcgdcrgagppdlsrlsvltvtvnllepldedlfrgrhywvpakrlfsggqivgqalvaa 69

Qy 62 SVSEDDVHVSILCHYFVRAGDPKLPVLVQVVERTTSSFSVRSVKAVOH-CKPIETCO 117

Db 70 svseddvhvsilhcyfvrsgdpkplvlyqvverttsgsfsvrskvavoh-ckpficq 129

Qy 118 ASFQQAQPPSMQHOFMSMPTVPPPELLDCEITLIDQYLDPNLQKRYPLALNRIAAQGV 175

Db 130 asfqaqapgmghfsmptvpppeellidcetlidyldpnlqkryplalnriaaqev 189

Qy 176 EIEIKVPNPSKL-SQQLRMEPKMFWRNARGYIGEG-DKMKHCVCVAYISDYAFELGTALL 233

Db 190 pieikvnpnpklrmeepkqmfwrargyigkspkwmkhcvcvayisdyafigtalla 249

Qy 234 P-HQW-OHKRVHFM-VSLDHS-NWGHAPRADHMLIYECESPMWAGSGRLVHGLWRQDGV 289

Db 250 altvqgstchrfhuppldnpmwfhagfradhwlmlyecespmwagsrglvhglwrqdg 309

Qy 290 LAVTCAQEGVIRVKPOVSESkl 311

Db 310 lavtcaqevlrvkqvqvseskl 331

PA (CORR) CORNELL RES FOUND INC.
 PI Yoder OC, Turgeon BC, Lu S;
 DR WPI: 2001-367672/38.
 XX New isolated nucleic acid molecule from a plant pathogen useful in
 PT preventing plant pathogenic infections -
 XX disclosure; Fig 22; 132pp; English.
 XX The sequence represents a thioesterase type II consensus sequence. These
 CC thioesterases include Cochliobolus heterostrophus TESI, thioesterase I,
 CC the gene for which is located in the same genomic region as the peptide
 CC synthetase, CPSI. CPSI is an enzyme thought to be involved in the
 CC production of peptide toxins, which are involved in the pathogenic
 CC infection of corn crops. The nucleic acids and proteins can be used as
 CC targets for anti-fungal compounds to prevent fungal corn infection and
 CC the nucleic acids can be used in gene therapy to alter the biosynthetic
 CC pathway for the peptide toxins to lower the pathogenicity of the fungi.
 XX Sequence 270 AA;
 XX
 Query Match 37.7%; Score 628.5; DB 22; Length 270;
 Best Local Similarity 44.2%; Pred No. 2, 1e-61;
 Matches 138; Conservative 11; Mismatches 78; Indels 85; Gaps 4;
 OY 25 NLEPLDRLFRGRHWVP---AKRLFGQIVGQALVAAKSVSDVHSHLCYFVRAGD 81
 DB 4 dleplexifgxxxxxxgarrxfggvgvgalvaaxktvdxvshlcyflragd 63
 OY 82 PKLPLVLYQVETRGSSFSVRSVAVOHGKPIFCOASFQQAQSPMQHOFSPMPTVPPE 141
 DB 64 skxpvlyxvrrxgxfstrrvavqhgkpfixmtasfgxxqgxxqhxmpxap--- 120
 OY 142 ELLDCETLIDQVLRDPLNOKRYPLALNRITAAQEVPIEIKVPWPSLSQLQRMPEKQFW 201
 DB 121 -----xxxxxepqgvwx 133
 OY 202 RARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHVF----- 243
 DB 134 targsxxpgdxxhxcalaymsdixfignlqphxlxxxxxxkxxxxxxkxxxx 193
 OY 244 -----MVLSDHSMWFH--APFRADHWMLYECESPWAGSGLVHGRLWRQDGV 289
 DB 194 xxxxxdxxxxqmasidhwmfhrxxpfradewillyexesagggrglvrgxytdgv 253
 290 LAVTCAQEGVIR 301
 DB 254 lvascvqevxvr 265
 RESULT 7
 ID AAY08619 standard; Protein; 286 AA.
 XX AAY08619;
 XX
 XX 05-AUG-1999 (first entry)
 XX E. coli TEII protein.
 XX
 XX PxTE; peroxisomal thioesterase; treatment; disorder; biosynthesis;
 XX fatty acid metabolism; chain-termination; fatty acid; cancer; deficiency;
 XX inflammation; adrenoleukodystrophy Refsum's disease; mental retardation;
 XX disordered neuronal migration; hypotonia; tapetoretinal degeneration; TEII.
 XX sensorineural hearing loss; cystic change; kidney; skeletal change; TEII.
 XX Escherichia coli.
 OS
 XX
 XX US5911984-A.
 XX

PD 15-JUN-1999.
 XX 19-JUN-1998; 98US-0100851.
 XX 11-JUN-1997; 97US-0872784.
 XX 19-JUN-1998; 98US-0100851.
 XX (INCY-) INCYTE PHARM INC.
 XX Corley NC, Hillman JL, Shah P;
 PI WPI: 1999-357169/30.
 XX
 XX Peroxisomal thioesterase enzymes useful for treating disorders
 PT associated with fatty acid metabolism
 XX disclosure; Fig 2A-B; 29pp; English.
 XX This invention describes a novel human peroxisomal thioesterase enzyme
 CC (PxTE) and its use to treat disorders associated with fatty acid
 CC metabolism. Thioesterases catalyze the chain-terminating step in the
 CC de novo biosynthesis of fatty acids. PxTE is a novel thioesterase
 CC involved in disorders associated with fatty acid metabolism and the
 CC inflammation of PxTE-deficient patients associated with Refsum's disease, and
 CC PxTE deficiency. Patients suffering from defective peroxisomal fatty
 CC acid metabolism exhibit disorder of neuronal migration, hypotonia, mental
 CC retardation, tapetoretinal degeneration, sensorineural hearing loss,
 CC cystic changes in the kidney, skeletal changes and death. This sequence
 CC represents the Escherichia coli TEII protein which is used to illustrate
 CC the method of the invention
 XX Sequence 286 AA;
 XX
 Query Match 33.2%; Score 553; DB 20; Length 286;
 Best Local Similarity 42.1%; Pred. No. 6e-53;
 Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;
 OY 15 LRSVLVTTLNLEPLDRLFRGRHWVPKAKRLFGQIVGQALVAAKSVSDVHSHLCR 74
 DB 5 Lknll--tlilnleegifrggsdglrvgvgvgalvaaketyceerlvhsfhs 62
 OY 75 YEVKAGDKPLVLYQVETRGSSFSVRSVAVOHGKPIFCOASFQQAQSPMQHOFSM 134
 DB 63 yfLrpgdskpilydvdelrldgnsfsarrvvaalngkpfifmtasf-gapeagfeghktm 121
 OY 135 PTVPPPELLDCELTIDQ---YLKDPNLOKRYPLALNRITAAQEVPIEIKVPN-PSPLSQL 190
 DB 122 psapadg-lpsetqlagslahlppvlkdkf-----icdrplervpvefhnpklg- 171
 OY 191 QRMEPKOMFWVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHVF----- 244
 DB 172 hvaeprhqvwfrangsvpd-dirvhgyligyasdnflpvalqp---hgfgleegiqi 226
 OY 245 VSLDSHSMWFHAPFRADHWMLYECESPWAGSGLVHGRLWRQDGVAVTCAQEGVIR 301
 DB 227 atidshwmfhrpfnlnewillyvestsassaargfvrgfytqdgvlvastvqevgvmr 283
 RESULT 8
 ID AAG24648 standard; Protein; 271 AA.
 XX AAG24648;
 XX 17-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 28404.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.

OS Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0125786.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126288.
XX 29-MAR-1999; 99US-0126795.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132048.
XX 05-MAY-1999; 99US-0132407.
XX 06-MAY-1999; 99US-0132484.
XX 07-MAY-1999; 99US-0132485.
XX 11-MAY-1999; 99US-0132486.
XX 14-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 28-MAY-1999; 99US-0136392.
XX 01-JUN-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139494.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139464.
XX 21-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140354.
XX 28-JUN-1999; 99US-0140354.
XX 29-JUN-1999; 99US-0140354.
XX 99US-0141287.
XX 99US-0141842.
XX 99US-0142154.
XX 99US-0142055.
XX 99US-0142390.
XX 99US-0142803.
XX 99US-0142920.
XX 99US-0142977.
XX 99US-0143542.
XX 99US-0143624.
XX 99US-0144005.
XX 99US-0144085.
XX 99US-0144086.
XX 99US-0144325.
XX 99US-0144331.
XX 99US-0144332.
XX 99US-0144333.
XX 99US-0144335.
XX 99US-0144335.
XX 99US-0144352.
XX 99US-0144632.
XX 99US-0144884.
XX 99US-0144814.
XX 99US-0145086.
XX 99US-0145088.
XX 99US-0145085.
XX 99US-0145087.
XX 99US-0145089.
XX 99US-0145192.
XX 99US-0145145.
XX 99US-0145218.
XX 99US-0145224.
XX 99US-0145276.
XX 99US-0145913.
XX 99US-0145918.
XX 99US-0145919.
XX 99US-0145951.
XX 99US-0146386.
XX 99US-0146388.
XX 99US-0146389.
XX 99US-0147038.
XX 99US-0147204.
XX 99US-0147302.
XX 99US-0147192.
XX 99US-0147260.
XX 99US-0147303.
XX 99US-0147416.
XX 99US-0147493.
XX 99US-0147935.
XX 99US-0148171.
XX 99US-0148319.
XX 99US-0148341.
XX 99US-0148565.
XX 99US-0148684.
XX 99US-0149368.
XX 99US-0149175.
XX 99US-0149426.
XX 99US-0149722.
XX 99US-0149723.
XX 99US-0149929.
XX 99US-0149929.
XX 99US-0149930.
XX 99US-0149930.
XX 99US-0150566.
XX 99US-0150884.
XX 99US-0151085.
XX 99US-0151086.
XX 99US-0151080.
XX 99US-0151303.
XX 99US-0151348.
XX 01-SEP-1999; 99US-0151940.
XX 07-SEP-1999; 99US-0152363.
XX 10-SEP-1999; 99US-0153070.
XX 13-SEP-1999; 99US-0153758.

```

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159339.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160757.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160816.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.4%; Score 489; DB 21; Length 271;
Best Local Similarity 38.8%; Pred. No. 7.7e-46;
Matches 101; Conservative 52; Mismatches 93; Indels 14; Gaps 4;

14 DLRSVLVTVTLNLEPDEDLGRHYWVPAK-----RLRGQIVGQALVAAKSVSDVH 68
17 dkerclvedilylepdlhnyrg---ftpnaptkygvygglygqalaasaklvetmk1 73

69 VHSLHCYFVRAGDPKPLVLYQVETRIGTSFSVRSKAVOHGKPIFICQASFOQAQSPM 128
74 vhmthcyfllygdiniplyadvnrlrdgnntatrsdvarqgkttiflfasfgkkggfi 133

129 OHQFSMPTVPPEELDCETLDQYLRDNLQKRYPLALNRIAAQ---EVPTEIKPVNPS 185
134 hqgestmhpcepelcpreemlerlvtelprdy---rnqvateisvfpdlrfcepn 190
186 PLSGLQRMPEKQMFWRVARYIGEGDKMKHCCVAAVYLSDYAFPLGTALLPHQNHKHYFW 245
191 rstqknspprikywfrakgksldddqalhrccvafasdlifatislnphrrfegmsvaal 250

246 SLDHSHWAFHFRADHWMLY 265
251 sldhsmwfrhfrvradgwllf 270

RESULT 9
AAG24647
ID AAG24647 standard; Protein: 377 AA.
XX
AC AAG24647;

```

```

XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 28403.
DE Protein identification; signal transduction pathway: metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127652.
XX 06-APR-1999; 99US-0128734.
XX 08-APR-1999; 99US-0128744.
XX 10-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138647.
XX 16-JUN-1999; 99US-0139119.
XX 17-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.

```

```
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 26-JUN-1999; 99US-0140991.
PR 26-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142152.
PR 01-JUL-1999; 99US-0142355.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147262.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147309.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148071.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149529.
PR 23-AUG-1999; 99US-0149502.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157425.
PR 02-OCT-1999; 99US-0158093.
PR 07-OCT-1999; 99US-0158423.
PR 08-OCT-1999; 99US-0158426.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.4%; Score 489; DB 21; Length 377;
Best Local Similarity 38.4%; Pred. No. 1.2e-45;
Matches 101; Conservative 52; Mismatches 93; Indels 14; Gaps 4;

OY 14 DLRSLVTTVLNLEPLDLDLFRGRHYWPAK-----RLFGQIVGQALVAASKVSEDVH 68
Db 123 dkerclivedilylepdlinvryg---ftppnptygykvgvgqilvgalaaaskvetnki 179
OY 69 VUSLCTFVRAGDPKPLVLYQVVERTTCSSFSVRSVKAVQHGKPIFICQASFQQAQSPM 128
Db 180 vnmfbcyflilvgdinpiydvnmrlrdgnafatrsvarqkgktlftifafskkqgffi 239
OY 129 QHQFSMTVPVPPPELLCCETLIDQYLRDPNLQKRYPLALNRIAAQ---EYPIEKVPNPS 185
Db 240 hqesumtpatpetlpreemlierivteplprdy---rnqvateisvptldirfcep 296
OY 186 PLSQLORMPEKQMFVWRYGTCBGDMKHCVCVAAYISDVAFLGTALPHQWQHVKHFMV 245
Db 297 rctkqkspbprikywfrakgklsdddqalhrvcvafasdlifatisinphrregmsvaal 356
OY 246 SLDSHMFHAPFADHNMVLY 265
```

us-09-766-366-1.rag

Db					357	stohsmwfrhprvradwili	376
RESULT 10							
ACG24646							
XX	ACG24646 standard; Protein; 385 AA.						
XX							
XX	ACG24646;						
XX							
DT	17-OCT-2000 (first entry)						
DE	XX						
DE	XX Arabidopsis thaliana protein fragment SEQ ID NO: 28402.						
KW	Protein identification; signal transduction pathway; metabolic pathway;						
KW	hybridisation assay; genetic mapping; gene expression control; promoter;						
KW	termination sequence.						
OS	Arabidopsis thaliana.						
EP	EPI033405-A2.						
PD	06-SEP-2000.						
XX							
XX	25-FEB-2000; 2000EP-0301439.						
PR	25-FEB-1999; 9805-0121825.						
PR	05-MAR-1999; 9805-0123180.						
PR	09-MAR-1999; 9805-0123548.						
PR	23-MAR-1999; 9805-0125788.						
PR	25-MAR-1999; 9805-0126264.						
PR	29-MAR-1999; 9805-0126785.						
PR	01-APR-1999; 9805-0127462.						
PR	06-APR-1999; 9805-0128234.						
PR	08-APR-1999; 9805-0128714.						
PR	16-APR-1999; 9805-0129845.						
PR	19-APR-1999; 9805-0130077.						
PR	21-APR-1999; 9805-0130449.						
PR	23-APR-1999; 9805-0130510.						
PR	23-APR-1999; 9805-0130891.						
PR	28-APR-1999; 9805-0131449.						
PR	30-APR-1999; 9805-0132048.						
PR	30-APR-1999; 9805-0132407.						
PR	04-MAY-1999; 9805-0132484.						
PR	05-MAY-1999; 9805-0132485.						
PR	06-MAY-1999; 9805-0132486.						
PR	06-MAY-1999; 9805-0132487.						
PR	17-MAY-1999; 9805-0132863.						
PR	11-MAY-1999; 9805-0134256.						
PR	14-MAY-1999; 9805-0134216.						
PR	14-MAY-1999; 9805-0134223.						
PR	14-MAY-1999; 9805-0134370.						
PR	18-MAY-1999; 9805-0134768.						


```
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149922.
PR 23-AUG-1999; 99US-0149922.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-01511303.
PR 31-AUG-1999; 99US-01511438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.4%; Score 489; DB 21; Length 385;
Best Local Similarity 38.8%; Pred. No. 1.3e-45;
Matches 101; Conservativity 52; Mismatches 93; Indels 14; Gaps 4;

QY 14 DURSVLVTYVLMLEPDEDFRGRHYWVPK-----RUGGQIVGQALVAAKSVSDVH 68
DB 131 kdercivdilylepidlnvryr---fupnaptkygkyvggqlvgqalaastketvtnki 187
QY 69 VHSIHCTFVRAGDKPLVLYQVETFRFGSSFSVRSKAVOHGKPTIFICQAFQQAQSPM 128
DB 188 vhnficyfllvgdlnpldyvnrldgnnfatrsvdardkgkgtlftifasfkgkqggfi 247
```

```
QY 129 QHOFNMTVPPEELDCETLIDYLRDPNLOKRYPLALNRAAQ---EVPETKVPNS 185
DB 248 hgestmptipatelilpreemilrvteplirpy---rnqvateisvppfidirfcpn 304
QY 186 PLSQJORMEPKQMFVVRARGYIGEGDMKHCVCVAAYISDYAFGLTALPQHMQKHVFW 245
DB 305 rsccknqepplkywfrakglisddqalhrvcvafasdlifatislnphregmsvaal 364
QY 246 SLDRSMWFHAPFRADHWMY 265
DB 365 slhsmwfhprvradgwllf 384

RESULT 11
ABG20246
TI ABG20246 standard; Protein; 106 AA.
AC ABG20246;
DT 13-FEB-2002 (first entry)
XX Human human diagnostic protein #20337.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX W0200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
DR N-PSDB; AAS84433.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 50605; 103pp; English.
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: the sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
```

SQ Sequence 106 AA;

Query Match 29.1%; Score 484.5; DB 22; Length 106;
 Best Local Similarity 60.8%; Pred. No. 6.3e-46;
 Matches 101; Conservative 0; Mismatches 0; Indels 65; Gaps 1;

QY 146 CETLIDQLRPNLQKRYPLALNRIAAQGVETIKPNVSPUSQLQRMPEKQMFWRARG 205
 D 6 cctldqlrpnalqkryplalnriaaqgevpelkpnvpsplslqrmepkqmtwvzarg 65
 QY 206 YIGEGDKMKCCVAAAYISDYAFGLTALLPHQMHQKVFHMYSLDHSMMFHAFRADHWMLY 265
 D 66 yi----- 67

QY 266 ECESPWAGSGRLVHGLRWRODGVLAATCAQGVIRVKPQVSSEKL 311
 D 68 -----gsgslvghrlwrqdgvlavtcaqgevirvkpqvsesk1 106

BLAST 12

91582
 ID: AAG91582 standard; Protein: 282 AA.

XX AAG91582;

XX AAG91582;

DT 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 5336.
 DE

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.

XX Corynebacterium glutamicum.
 OS

XX EP1108790-A2.
 PN

XX 20-JUN-2001.
 PD

XX 18-DEC-2000; 2000EP-0127688.
 PF

XX 16-DEC-1999; 99JP-0377484.
 PR

XX 07-APR-2000; 2000JP-0159162.
 PR

XX 03-AUG-2000; 2000JP-0280988.
 PR

XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochial K, Yokoi H;
 PT Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.
 DR N-PSDB; AAH56801.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT

XX Claim 17; SEQ ID NO: 5336; 246pp + Sequence Listing; English.
 PS

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX SQ Sequence 282 AA;

Query Match 28.0%; Score 466.5; DB 22; Length 282;
 Best Local Similarity 36.6%; Pred. No. 2.6e-43;
 Matches 107; Conservative 45; Mismatches 119; Indels 21; Gaps 7;

QY 20 VTTVNLNLEPLDEDLFRGRHYWVPKAKLFGGQIVGQALVAAAKSVSDVHVHSLHCYFVRA 79
 D 4 iedllleldrdldyrgpviessylartfggvaqaalvaatntvdkaftvhaihgylfep 63
 D 4 iedllleldrdldyrgpviessylartfggvaqaalvaatntvdkaftvhaihgylfep 63

QY 80 GDPKPLVLYQVVERFRGSGFSVRSKAVQHGKPIFCQASFOQAQSPMQHQFSMPTVPP 139
 D 64 gdpkplvlydvrvdrgskvrsrglqdgdevlfsmqasfhrdeg-lehmdkmrvkpa 122
 D 64 gdpkplvlydvrvdrgskvrsrglqdgdevlfsmqasfhrdeg-lehmdkmrvkpa 122

QY 140 PBEELDCETLIDQLRDPNLOKRYPLALNRIAAQ--EVPTEIKPNVSPUSQLQRMPEKQ 197
 D 123 pdei-----kgtvermpisrrvldewaewdirvipqdgelsafate-q 167

QY 198 MFVVRARGYIGEGDKMKHCCVAAAYISDYAFGLTALLPHQMHQKVFHMYSLDHSMMFHAF 257
 D 168 avlwrctadlpl-nptfhqcsaltylemtlinalvphqge-kmq-masldhavvlrpf 224

QY 258 RADHWMLYECESPWAGSGRLVHGLRWRODGVLAATCAQGVIRVKPQVSSES 309
 D 225 rvdewillydqrspsaasgralthrgrlfnqgdlvalvndegmtrtlhdegaqs 276

RESULT 13

AAAB76590
 ID: AAB76590 standard; Protein: 282 AA.

XX AAB76590;

XX AAB76590;

DT 11-APR-2001 (first entry)

XX Corynebacterium glutamicum MCT protein SEQ ID NO:162.
 DE

XX Corynebacterium glutamicum; brevbacterium lactofermentum; MCT;
 KW membrane construction and membrane transport protein; petroleum spill;
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
 KW identification; microorganism; fine chemical production; transformation;
 KW genome mapping; genetic engineering.

XX Corynebacterium glutamicum.
 OS

XX WO200100805-A2.
 PN

XX 04-JAN-2001.
 PD

XX 23-JUN-2000; 2000WO-IB00926.
 PF

XX 25-JUN-1999; 99US-0141031.
 PR

XX 08-JUL-1999; 99DE-1031454.
 PR

XX 08-JUL-1999; 99DE-1031478.
 PR

XX 08-JUL-1999; 99DE-1031563.
 PR

XX 09-JUL-1999; 99DE-1032122.
 PR

XX 09-JUL-1999; 99DE-1032124.
 PR

XX 09-JUL-1999; 99DE-1032125.
 PR

XX 09-JUL-1999; 99DE-1032128.
 PR

XX 09-JUL-1999; 99DE-1032180.
 PR

XX 09-JUL-1999; 99DE-1032190.
 PR

XX 09-JUL-1999; 99DE-1032209.
 PR

XX 09-JUL-1999; 99DE-1032212.
 PR

XX 09-JUL-1999; 99DE-1032227.
 PR

XX 09-JUL-1999; 99DE-1032248.
 PR

XX 09-JUL-1999; 99DE-1032249.
 PR

XX 14-JUL-1999; 99DE-1032250.
 PR

XX 14-JUL-1999; 99DE-1032927.
 PR

XX 14-JUL-1999; 99DE-1033005.
 PR

PR 14-JUL-1999; 99DE-1033006.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040830.
PR 27-AUG-1999; 99DE-1040831.
PR 27-AUG-1999; 99DE-1040832.
PR 27-AUG-1999; 99DE-1040833.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042078.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042088.
PA (BADI) BASF AG.
XX
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI: 2001-071486/08.
DR N-PSDB; AAF67823.
XX
PT Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation -
XX
XX Claim 20; Page 392-393; 1119pp; English.
XX
CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAF67510 to
CC AAF67847. The MCT proteins and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.
XX
XX Sequence 282 AA:
XX
XX
XX Query Match 28.0%; Score 466.5; DB 22; Length 282;
XX Best Local Similarity 36.6%; Pred. No. 2.6e-43;
XX Matches 107; Conservative 45; Mismatches 119; Indels 21; Gaps 7;
XX
XX 20 VTTVLNLEPLEDELFGRHHVWPAAKRLFGQIYGGALVAAKSVSEDEVHSLHCYFVRA 79
XX D 4 iedltleedrdygrgvyiesygartfgrgvaalvaathvdkatfshslyhyflap 63
XX
XX 80 GPKLPLVLYOVERTTSSSVRSKAVQHGKPTFCQASFOQAQSPMQHQFSMTPTVP 139
XX Db 64 gdpdpalyldvrvdrgskysvtrsvrgldgdevifsmqasfhrddeg-iehmkmrkypa 122
XX
XX 140 PELLDCETLIDQYLEDNDNQKRYPLALNIAAQ--EVPTEIKPVNPSLSQLORMPEKQ 197
XX Db 123 pdei-----kgvempisrrvldewaedwdrvipdqqlsdfate--q 167
XX
XX 198 MWKRVANGYIGCEDMKHCYAAVSYAFYALLPHQHQHVHNVSLDHSMWFHAPF 257
XX Db 168 awvictadltp nptfhqesityisdmtllsalvnpge-kmq-masidhavwlrpf 244
XX
XX 258 RADHWMLYECESPNAGSGRLVHGRLHRODGLVATCAOBCVIRVKPQVSES 309
XX Db 225 rvdewllydgrspasagralthrlfnqgddlvaivnqegmcrllthegaaqs 276
XX
XX
XX RESULT 14
XX AAF67591
XX ID AAF67591 standard; Protein; 282 AA.
XX

AC AAF67591;
XX
XX 11-APR-2001 (first entry)
XX
XX Corynebacterium glutamicum MCT protein SEQ ID NO:164.
XX
XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
XX membrane construction and membrane transport protein; petroleum spill;
XX hydrocarbon degradation; gram positive aerobic bacterium; marker;
XX identification; microorganism; fine chemical production; transformation;
XX genome mapping; genetic engineering.
XX
XX Corynebacterium glutamicum.
XX
XX WQ200100805-A2.
XX
XX 04-JAN-2001.
XX
XX
XX 23-JUN-2000; 2000WO-IB00926.
XX
XX 25-JUN-1999; 99US-0141031.
XX 08-JUL-1999; 99DE-1031454.
XX 08-JUL-1999; 99DE-1031478.
XX 08-JUL-1999; 99DE-1031563.
XX 09-JUL-1999; 99DE-1032122.
XX 09-JUL-1999; 99DE-1032124.
XX 09-JUL-1999; 99DE-1032125.
XX 09-JUL-1999; 99DE-1032128.
XX 09-JUL-1999; 99DE-1032180.
XX 09-JUL-1999; 99DE-1032182.
XX 09-JUL-1999; 99DE-1032190.
XX 09-JUL-1999; 99DE-1032209.
XX 09-JUL-1999; 99DE-1032212.
XX 09-JUL-1999; 99DE-1032227.
XX 09-JUL-1999; 99DE-1032228.
XX 09-JUL-1999; 99DE-1032229.
XX 09-JUL-1999; 99DE-1032230.
XX 14-JUL-1999; 99DE-1032927.
XX 14-JUL-1999; 99DE-1033005.
XX 14-JUL-1999; 99DE-1033006.
XX 27-AUG-1999; 99DE-1040764.
XX 27-AUG-1999; 99DE-1040765.
XX 27-AUG-1999; 99DE-1040766.
XX 27-AUG-1999; 99DE-1040830.
XX 27-AUG-1999; 99DE-1040831.
XX 27-AUG-1999; 99DE-1040832.
XX 27-AUG-1999; 99DE-1040833.
XX 31-AUG-1999; 99DE-1041378.
XX 31-AUG-1999; 99DE-1041379.
XX 31-AUG-1999; 99DE-1041395.
XX 03-SEP-1999; 99DE-1042077.
XX 03-SEP-1999; 99DE-1042078.
XX 03-SEP-1999; 99DE-1042079.
XX 03-SEP-1999; 99DE-1042088.
XX
XX (BADI) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI: 2001-071486/08.
XX N-PSDB; AAF67824.
XX
XX Corynebacterium glutamicum nucleic acids encoding membrane construction
XX and membrane transport proteins or their portions, useful for typing or
XX identifying C. glutamicum or related bacteria, and as markers for
XX transformation -
XX
XX Claim 20; Page 395-396; 1119pp; English.
XX
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
XX construction and membrane transport (MCT) proteins given in AAF67510 to
XX AAF67847. The MCT proteins and proteins are useful in the

GenCode version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 13:18:29 ; Search time 13.12 Seconds
(without alignments)
578.991 Million cell updates/sec

Title: US-09-766-366-1
Perfect score: 1666
Sequence: 1 MSRAVATAALPPGDLRSVLV.....VTCAQGVIRKPVQSVBSKL 311

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.psp.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.psp.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.psp.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.psp.*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.psp.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.psp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1666	100.0	311	1	US-08-872-784-1
2	1666	100.0	311	2	Sequence 1, Appl
3	1666	100.0	311	4	US-09-100-851-1
4	553	33.2	286	1	US-08-872-784-1
5	553	33.2	286	2	Sequence 3, Appl
6	553	33.2	286	4	US-09-265-294-3
7	273	16.4	349	1	US-08-872-784-1
8	273	16.4	349	2	Sequence 4, Appl
9	273	16.4	349	4	US-09-100-851-4
10	100	6.0	1580	2	US-08-804-227C-11
11	100	6.0	1580	2	US-08-804-198-5
12	89	5.3	2325	3	US-08-417-089-6
13	89	5.3	2325	4	US-08-695-651-6
14	89	5.3	2325	4	US-08-930-285-6
15	89	5.3	2325	4	US-08-695-421-6
16	83.5	5.0	529	5	PCT-0592-00282-7
17	80.5	4.8	516	4	US-08-867-611-6
18	80.5	4.8	516	5	PCT-0592-06969A-11
19	80.5	4.8	798	4	US-08-867-611-36
20	79.5	4.8	1068	4	US-09-085-199B-11
21	79	4.7	412	4	US-09-461-637-77
22	78.5	4.7	525	1	US-08-160-861-4
23	78.5	4.7	680	4	US-08-342-363-4
24	78.5	4.7	680	4	US-09-100-889-4
25	78.5	4.7	1350	4	US-08-319-866-9
26	78	4.7	481	3	US-08-789-311-2
27	77.5	4.7	732	1	US-08-317-522A-5

ALIGNMENTS

RESULT 1
US-08-872-784-1
: Sequence 1, Application US/08872784
: Patent No. 5776753
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/872,784
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0293 US
: TELECOMMUNICATIONS INFORMATION:
: TELEPHONE: 415-845-0555
: FAX: 415-845-2166
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: TYPE: amino acid
: LENGTH: 311 amino acids
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: BRAINNOT09
: CLONE: 2150905
US-08-872-784-1

Query Match 100.0% Score 1666; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e-182;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 24, Appl
Sequence 4, Appl
Sequence 9, Appl
Sequence 26, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 9, Appl
Sequence 2, Appl

QY 1 MGRAVATAALPGDLRSVLVTTVNLNLEPLDEDLFRGRHYWVPAKRLFGGOIVGOALVAAA 60
 DB 1 MGRAVATAALPGDLRSVLVTTVNLNLEPLDEDLFRGRHYWVPAKRLFGGOIVGOALVAAA 60
 QY 61 KSVSDEVHSHLCYFVRAGDPKLPVLYQVTRTGTSSFSVRSVKAQVHGKPIFTCOASF 120
 DB 61 KSVSDEVHSHLCYFVRAGDPKLPVLYQVTRTGTSSFSVRSVKAQVHGKPIFTCOASF 120
 QY 121 QQAQSPMHOQSMPTVPPPELLDCETLIDQYLRDPNLOKRYPLALNRIAAOEVPPIEK 180
 DB 121 QQAQSPMHOQSMPTVPPPELLDCETLIDQYLRDPNLOKRYPLALNRIAAOEVPPIEK 180
 QY 181 PVNPSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 240
 DB 181 PVNPSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 240
 QY 241 VHPVMSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 300
 DB 241 VHPVMSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 300
 QY 301 RVKQVSESCL 311
 DB 301 RVKQVSESCL 311

RESULT 2
 : US-09-100-851-1
 : Sequence 1, Application US/09100851
 : Patent No. 5911984
 : GENERAL INFORMATION:
 : APPLICANT: Hillman, Jennifer L.
 : APPLICANT: Shah, Purvi
 : APPLICANT: Corley, Neil C.
 : TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/100,851
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/872,784
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J.
 : REGISTRATION NUMBER: 36,749
 : REFERENCE/DOCKET NUMBER: PF-0293 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415-855-0555
 : TELEFAX: 415-845-4166
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 311 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: BRAINOT09
 : CLONE: 2150905

US-09-100-851-1

Query Match 100.0%; Score 1666; DB 2; Length 311;
 Best Local Similarity 100.0%; Pctd No. 1, 5e-182; Indels 0; Gaps 0;
 Matches 311; Conservative 0; Mismatches 0;
 QY 1 MGRAVATAALPGDLRSVLVTTVNLNLEPLDEDLFRGRHYWVPAKRLFGGOIVGOALVAAA 60
 DB 1 MGRAVATAALPGDLRSVLVTTVNLNLEPLDEDLFRGRHYWVPAKRLFGGOIVGOALVAAA 60
 QY 61 KSVSDEVHSHLCYFVRAGDPKLPVLYQVTRTGTSSFSVRSVKAQVHGKPIFTCOASF 120
 DB 61 KSVSDEVHSHLCYFVRAGDPKLPVLYQVTRTGTSSFSVRSVKAQVHGKPIFTCOASF 120
 QY 121 QQAQSPMHOQSMPTVPPPELLDCETLIDQYLRDPNLOKRYPLALNRIAAOEVPPIEK 180
 DB 121 QQAQSPMHOQSMPTVPPPELLDCETLIDQYLRDPNLOKRYPLALNRIAAOEVPPIEK 180
 QY 181 PVNPSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 240
 DB 181 PVNPSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 240
 QY 241 VHPVMSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 300
 DB 241 VHPVMSLSQRMPEKQMFVRARGYIGEGDMKHCVCVAAVYISDYAFGLTALLPHQWQHK 300
 QY 301 RVKQVSESCL 311
 DB 301 RVKQVSESCL 311
 RESULT 3
 : US-09-265-294-1
 : Sequence 1, Application US/09265294
 : Patent No. 6210890
 : GENERAL INFORMATION:
 : APPLICANT: Hillman, Jennifer L.
 : APPLICANT: Shah, Purvi
 : APPLICANT: Corley, Neil C.
 : TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/265,294
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 09/100,851
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J.
 : REGISTRATION NUMBER: 36,749
 : REFERENCE/DOCKET NUMBER: PF-0293 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415-855-0555
 : TELEFAX: 415-845-4166
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 311 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: BRAINOT09
 : CLONE: 2150905

US-09-265-294-1

```

Query Match      100.0%; Score 1666; DB 4; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e-182;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRVATAALPPGDLRSVVTVNLNLEPDELDFGRHYWVPAKRLGGQVGOALVAAA 60
DQ 1 MGRVATAALPPGDLRSVVTVNLNLEPDELDFGRHYWVPAKRLGGQVGOALVAAA 60
QY 61 KVSVDSEVHSLHCVFVYRAGDKPLVLYQVETRTGSGFSVRSVKAQVHGKPIFICQASF 120
DQ 61 KVSVDSEVHSLHCVFVYRAGDKPLVLYQVETRTGSGFSVRSVKAQVHGKPIFICQASF 120
QY 121 QQAQSPWQHQFSMTVPPPEELLDCETLLDQYLDNPKQKRYPLALNRITAAQVPIEIK 180
DQ 121 QQAQSPWQHQFSMTVPPPEELLDCETLLDQYLDNPKQKRYPLALNRITAAQVPIEIK 180
QY 181 PNPSPFLSQIQMRKQFVYRAGYIGEGDMKMCVAAVYISDYAFGLTALLPHQOHK 240
DQ 181 PNPSPFLSQIQMRKQFVYRAGYIGEGDMKMCVAAVYISDYAFGLTALLPHQOHK 240
QY 241 VHKWYSLDSEWFAHFRADHNMVYECESPWAGSGRGLVHGRWODQVGLAVTCAQEGVI 300
DQ 241 VHKWYSLDSEWFAHFRADHNMVYECESPWAGSGRGLVHGRWODQVGLAVTCAQEGVI 300
QY 301 RYKQVQSESKL 311
DQ 301 RYKQVQSESKL 311

```

RESULT 4

```

US-08-872-784-3
: Sequence 3, Application US/08872784
: Patent No. 5776753
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Shah, Purvi
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/872,784
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0293 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-853-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 286 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```

```

: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 147932
: US-08-872-784-3

```

```

Query Match      33.2%; Score 553; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 4.2e-55;
Matches 125; Conservative 53; Mismatches 91; Indels 28; Gaps 10;

QY 15 LRSVTVTVMLPELDEDFGRHYWVPAKRLGGQVGOALVAAKASVSDVHVSRLC 74
DQ 5 LKNNL--TLLNLEKIEGLRQSGEDGLRQVFGQVGOALVAAKAVPEERLVSEHS 62
QY 75 YFVRAQDKLPLVLYQVETRTGSGFSVRSVKAQVHGKPIFICQASFQQAQSPWQHFQSM 134
DQ 63 YFLRQDSKKPIYVETLRQNSFSARRVAIQNGKPIFTYMTASF-QAPEAGFEHOKTM 121
QY 135 YTPVPEELLDCETLLDQYLDNPKQKRYPLALNRITAAQVPIETKPNV-PSPLSOL 190
DQ 122 SFAAPDG-LPSETQIAQSLAHLPPVLKDKF-----ICDRPLEVHPVEPHNPLKG- 171
QY 191 ORMPKQFVYRAGYIGEGDMKMCVAAVYISDYAFGLTALLPHQOHKVFHVM----- 244
DQ 172 VHAEPHROWLRANGSVDP-DLRVHOYLLGYASDLNPLPALQV----HGIGFLEPGIQI 226
QY 245 VSLDSEWFAHFRADHNMVYECESPWAGSGRGLVHGRWODQVGLAVTCAQEGVIR 301
DQ 227 ATIDHSMFHPFNLEMLLYSVESSTASSARGFVGEFTODGVLVASTVQECVMR 283

```

RESULT 5

```

US-09-100-851-3
: Sequence 3, Application US/09100851
: Patent No. 5911984
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Shah, Purvi
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/100,851
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/872,784
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0293 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-853-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 286 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:

```

```

; LIBRARY: GenBank
; CLONE: 147932
US-09-100-851-3

```

Query Match 33.28; Score 553; DB 2; Length 286;

Qy	15	LRSLVTVTVMLEPDLDDLPGRHWYWPAPAKRLFGQITVGQALVAAAKSVSDVHVSHLHC	74
Db	5	LKWL--TLLMKLEEGFLQGSGDGLGRVFGQVVGQALYAAKATPEERLVSFHS	62
Qy	75	YFVRAGDKPLVLYQVERITGTGSSFSVRSKAVQGRKPIFCIQASFOQAQPSMQHQFSM	134
Db	63	YFVRGDSKPIIYDVTURDNGSFARVAAIQNGKPIFYATSP-QAPEAGFQHMT	121
Qy	135	PTVPPPEELLDCETLIDQ---YLDRDNLQKRYPLALNRITAAQGEVPIIKVNPSPLSQL	190
Db	122	PSAPADG-LPSETQTAQSLAHLPLPVLDKPF-----ICDRLEVRVEFEHNLKMG-	171
Qy	191	QRMEPKQFWVRARGYIGEGDMKNHCVAATISDYAFGLTALLPHQHQHVFHM-----	244
Db	172	HVAEPIHQVITRANGSVDP-DLRVHOYLLGLASDLNLPVALQP---HGIGFLPFGIQI	226
Qy	245	VSLDSHSMFHPAPRADHMNYECESYQWAGSGRGLVHGRWRQDGLVATCAQEGVIR	301
Db	227	ATTIDSHMFIRPENLNMWLYSVESTSASSARGVEGEVTDQGLVASTQVEGVR	283

RESULT	6
--------	---

```

US-09-265-294-3
? Sequence 3, Application US/09265294
? Patent No. 6210890
? GENERAL INFORMATION:
? APPLICANT: Hillman, Jennifer L.
? APPLICANT: Shah, Purvi
? APPLICANT: Corley, Neil C.
? TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Incyte Pharmaceuticals, Inc.
? STREET: 171 Porter Drive
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/265,294
? FILING DATE:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/100,851
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Billings, Lucy J.
? REGISTRATION NUMBER: 36,749
? REFERENCE/DOCKET NUMBER: PP-0293 US
? TELECOMMUNICATIONS INFORMATION:
? TELEPHONE: 415-845-0855
? TELEFAX: 415-845-1166
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 286 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? LIBRARY: GenBank

```

; CLONE: 147932
US-09-265-294-3

Query Match	33.7%	Score 553;	DB 4;	Length 286;
Best Local Similarity	42.1%	Word No. 4 2e-55;		
Matches	125;	Conservative 53;	Mismatches 91;	Indels 28; Gaps 10;
QY	15	LRSVLVTTLNLPLDDEDLGRGHVHWPAKRLFGGOIVGOALVAAKSVSDYVHHSLLHC	74	
DB	5	LKNLL--TLLNLEKTEEGLRGQSEDLCGLROVGGVGGQALYAARKETVPEEELVHSFHS	62	
QY	75	YFVRAGDPKLPVLVQVETRTGSSFSVRSVKVAHQGKPIFICQASFOQAQSPMQHQSM	134	
DB	63	YFLRGDSKPIIYDVELTLDGNSTSARRVAATONGKPIFYMTAGF-QAFPAFGHQKTM	121	
QY	135	VPVPPPELDCETLUDQ---YLDRPNLQKRYPLALNRIAAQGVETPIKPVN-PSPLSQL	190	
DB	122	PSNAPQD-UPSEQIQASLHLLPVLKQK---ICDRPLEVVRVPEHPLNG-171		
QY	191	QRMEPKQHWVYRQYTGECQDKMKHCVAAYTSDYFATGLTALLPHQHQHVEN-----	244	
DB	172	HVAEPHROVMIRANGSVPD-DLRVHQYLLQYASDLNLFVLPALQ---HCGIFLEPGTQI	226	
QY	245	VSLDSHMFHIAFPRADHWMLYECESPMAGGSKGLVHGLRWDRQDVLVATTCAGGVIR	301	
DB	227	ATIDSHMFHIFPNEMLLYSVETSSASRKGPRVGEFTYDQGVLAIVASTVOEGVNR	283	

RESULT 7

```

US-08-872-784-4
? Sequence 4, Application US/08872784
? Patent No. 5776753
? GENERAL INFORMATION:
? APPLICANT: Hillman, Jennifer L.
? APPLICANT: Shah, Purvi
? APPLICANT: Corley, Neil C.
? TITLE OF INVENTION: HUMAN PEROXISOMAL THIOESTERASE
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Incyte Pharmaceuticals, Inc.
? STREET: 3174 Porter Drive
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? OPERATING SYSTEM: IBM Compatible
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/872,784
? PRIORITY DATE: Filed Herewith
? APPLICATION DATA:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Billings, Lucy J
? REGISTRATION NUMBER: 36,749
? REFERENCE/DOCKET NUMBER: PF-0293 US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-855-0555
? TELEFAX: 415-845-4166
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 349 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? LIBRARY: GenBank
? CLONE: 854594

```


;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 854594
; US-09-265-294-4

Query Match 16.4% Score 273; DB 4; Length 349;
Best Local Similarity 26.2%; Pred. No. 6,7e-23;
Matches 86; Conservative 53; Mismatches 135; Indels 54; Gaps 9;
QY 23 VLNLLELDLDFGRHVVTPA---KRLFGGIVGQALVAAKSVSDVHVHSLHCYFV 77
DB 14 ILELVPLSTFVTK--YLPAAPVSGKFGTGLVQSULSLAHLVPLVFPFTSLHSYFI 71
QY 78 RAGCDPLVLQVVERTRGSSFSVRSKAVQHGKPIFCQASF--QQAQPSPMHQFSM 134
DB 72 KGDPRKTIYHVLNLRNGRNFHIKVSAYQNDKLTSMILFAVORSKHDSLQHWETI 131
135 PTV-----PPELDCETLIDQLRDPNQLQKRYPLALNRIAAQVPIETKPNVSPLSOL 190
DB 132 PQLGGQDPHRYEATSLFQKEVLDPQKLSFYASLSDRF--QDATSMKYVDAPQYGM 189
QY 191 QRMPEKQWFW-----VRARGYI-----GEGOMK 214
DB 190 EYQPKDMFYSAHTDELDFVKYRPPITVEHAGDESSLHWHHPYRIPKSTITPENDARY 249
QY 215 HCCVAAVYISDAFLGHTALLPHOMQKVVH--FMVSLDHSMFHAPFRADHMLYCECSPWAG 273
DB 250 NVAFAYLSDSLVTLLPIFYHNLPLYCHSEFSVSLDHTIYFHLQPLHVNMMLYLKISNPRSH 309
QY 274 GSRGLVHRLW--RQDGVLAATCAQEGVI 300
DB 310 WDKHLVQKGYDTQSGRIMASVQEGYV 337

RESULT 10
US-08-804-227C-11
; Sequence 11, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rosteck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) text only
; CURRENT APPLICATION DATA:
; FILING DATE: February 21, 1997
; CLASSIFICATION: C12N 1/24
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

;
; LENGTH: 1580 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-804-227C-11

Query Match 6.0% Score 100; DB 2; Length 1580;
Best Local Similarity 20.6%; Pred. No. 0.042;
Matches 66; Conservative 25; Mismatches 122; Indels 108; Gaps 11;
QY 28 FLDEDLFRGRHVVPAKRLFGGIVGQALVAAKSVSDVHVHSLHCYFVRAGDKPLVPL 87
DB 900 PLPTYAFSSKDRVWLPT-----AIGRRVAAPV-----DASAPGR 934
QY 88 IQVETRTGSSFSVRSKAVQHGKPIFCQAFQQAQPSPMHQFSMPTVPPE----- 141
DB 935 YRVWT-----PV-----ASDGRPSGRKLLVOTTGTAPEADTAAS 972
QY 142 ELLDCETLIDQLRDPNQLQKRYPLA--LNRIAAQ-----EVPTEIKPVN 183
DB 973 ALGAAQVYVERCLDPTTEAARVTLTERLAELDONPGLAGVLVLPGRPOSTAFANDSLD 1032
QY 184 PSLPSLOLQWME-----PKOMFWYRARGYIGEDMKHCCVAAVYISDYATGLTALLPHQW 237
DB 1033 PGTAAVLLVQVQDPAAPKARIWVYTRGAVGVGSEVPCAVGARGVGLGRVAALEVPVQW 1092
QY 238 QHKYHFWNSLDHSMFHAFRADHMLYCECSPHAGGR-----GLVHGR----- 282
DB 1093 GGLVDVAVGAGVREM-----RRVGVYAGGGEQVAVRGSGVFGRLVGVGV 1140
QY 283 ----LNRQDGVLAATCAQEGV 299
DB 1141 CGSCWRARGCVVYTGCLGSV 1161

RESULT 11
US-08-804-198-5
; Sequence 5, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rosteck, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P5113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3685
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1580 amino acids

```

; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-198-5

Query Match          6.0%; Score 100; DB 2; Length 1580;
Best Local Similarity 20.6%; Pred. No. 0.042;
Matches 66; Conservative 25; Mismatches 122; Indels 108; Gaps 11:

QY 28 PLDEDLFGRHYWPAKRLFGGQIVGQALVAAAKSVSDVHVHSLHCYFVRAGDKLPVL 87
DB 900 ELPTAFASDRYWLPT-----AICGRAVEAAPV-----DASAFGR 934

QY 88 YQVETRTGSSFSVRSVKAVOHGKPIFICQASFOAQSPMQHQSMTVPVPE----- 141
DB 935 YRVWT-----PV-----ASDSGRPSGRWLLVOTGTAPPEADTAAS 972

QY 142 ELIDCETLIDYLRDNLKRYPLA--LNRTAAQ-----EVPETKPVN 183
DB 973 ALGAAGVYVRECLDPTTEARVLTTERLAEADQPGLAGVLVPGRPSTAPADASPLD 1032

QY 184 FSPLSQORME-----PKMPFWRVARGYIGEGDMKMHCCVAAVYSDYAFIQTALLPHQW 237
DB 1033 PGTAAVLVYQAVPDAPKARLWVYTRGAVAGSGPCAVGARWGLRGVAALEVPQM 1092

QY 238 QHKVHFWSLDHSMFHAFPRADHMLYECSPWAGGSR-----GLVHOR----- 282
DB 1093 GGLVDVAVAGVREW-----RRVGVVYAGGEDQVAVRGGVFGRRLLVGVGR 1140

QY 283 ----LHQDQGLVATCAQEGV 299
DB 1141 GSGGVWRARGCVVYVIGLGV 1161

RESULT 12
US-08-417-089-6
; Sequence 6, Application US/08417089
; Patent No. 6045298
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
; TITLE OF INVENTION: OIL CONTENT OF PLANTS
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION NUMBER: US/08/417,089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-417-089-6

Query Match          5.3%; Score 89; DB 3; Length 2325;
Best Local Similarity 19.9%; Pred. No. 1.4;
Matches 57; Conservative 43; Mismatches 102; Indels 84; Gaps 13:

QY 26 LEFLDEDLFGRHYWPAKRLFGGQIVGQALVAAAKSVSDVHVHSLHCYFVRAGDKLP 85
DB 287 LDAIEPMYKACVTTTEEAASVACGVYPMIKASWGSGGKGRKHVH-----NDDEV 340

QY 86 VLYQVETRTGSSFSVRSVKAVOHGKPIFICQASFOAQSPMQHQSMTVPVPEELD 145
DB 341 ALF-----KQVQGVPGSPFLVIRLASQS-----RH-----LE 368
```

```

QY 146 CETLIDQY-----LRDPNLKRYPLALNRITAAQEVPIETKPVNPSLSQIORMEPKOM 198
DB 369 VOLLCDQYGNVAALHSRDCSVORRH-----OKI-TEEGPVTVAPRETVKALEQAA 418

QY 199 FWVRARGYIGEGDMKMHCCVAAVYSDYAFIQTALLPH-QMOHKV-----BFMV 245
DB 419 RLAKAVSYGVCAATVEY--LYSMETGDIYEL--ELNPLQVHEHPVTEWIAEVLNPAQAV 474

QY 246 SLDSHSMFHAFPRADHMLYECSPWAGSGRLVHGRLMRQDGVLA 291
DB 475 GMGTPLAQIPEIRRPYQMDY-----GGGYD-----IMRKTAAALA 508

RESULT 13
US-08-695-651-6
; Sequence 6, Application US/08695651
; Patent No. 6148667
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, B. G.
; APPLICANT: Somers, D. A.
; APPLICANT: Wyse, L.
; APPLICANT: Cronwald, J. W.
; APPLICANT: Egan, M. A.
; APPLICANT: Lutz, S.
; TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluch, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION NUMBER: US/08/695,651
; FILING DATE: 12-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/417089
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: 08/014326
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 07/917462
; FILING DATE: 21-JUL-1992
; APPLICATION NUMBER: 07/538674
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren, D
; REGISTRATION NUMBER: 06,440
; REFERENCE/DOC# NUMBER: 500,318US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3081
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: N-terminal
; FRAGMENT TYPE: N-terminal
US-08-695-651-6

Query Match          5.3%; Score 89; DB 4; Length 2325;
Best Local Similarity 19.9%; Pred. No. 1.4;
```

Matches 57; Conservative 43; Mismatches 102; Indels 84; Gaps 13;

```

QY 26 LEPDLEDFRGRHYVWPAKRLFGGOIVGOALVAAKSVSESDVHVSLHCYFVRAGDPKLP 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 LDAPEEMRYKACVTTTEAVASCVGVYFAMTKASWGGGKGKIRKVH-----NDDEV 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 86 VLYQVTRTGSFSFVSRSKAVOHGKPIFCQASFOQAPSPHQHOFSMPTVPPPELID 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 341 ALF-----KQVQGEVPGSPIFYMRLASQS-----RH-----LE 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 CETLIDQ-----LRDPLNOKRYPLALNRIAAQEVPIETKPVNPSLSQLQRMPEKPM 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 VOLLCDDQYGNVAALHSRDCSVQRH-----QKI-IEEGPVTVAPRETVKALEQAR 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 FVVRARYIGEGDMKHCVAAYISDYAFGLTALLPH-OMQHKV-----HPMV 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 RLAQAVGVGAATVEY--LYSMETGDYFVL--ELNPLQVEHPVTETAVNLPAAQAV 474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 246 SLDSHSMFHAPFRADHMLVCESPWAGGSRGLVHRLMRQDCVLA 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 475 GMGIPLQWQPEIRRFYGMDF-----GGGYD-----IWRKTAALA 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

```

US-08-930-285-6
: Sequence 6, Application US/08930285
: Patent No. 6221099
: GENERAL INFORMATION:
: APPLICANT: Regents of the University of Minnesota, et al.
: TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
: STREET: P. O. Box 2938
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/930,285
: FILING DATE: 13-APR-1998
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/04625
: FILING DATE: 04-APR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Emrictson, Janet E.
: REGISTRATION NUMBER: 39,665
: REFERENCE/DOCKET NUMBER: 600.318054
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-3031
: TELEFAX: 612-339-3061
: TELEX:

```

```

: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2325 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal and C-terminal (full length protein)
: ORIGINAL SOURCE:
: US-08-930-285-6

```

Query Match

5.3%; Score 89; DB 4; Length 2325;

Best Local Similarity 19.9%; Pred. No. 1.4; Matches 57; Conservative 43; Mismatches 102; Indels 84; Gaps 13;

```

QY 26 LEPDLEDFRGRHYVWPAKRLFGGOIVGOALVAAKSVSESDVHVSLHCYFVRAGDPKLP 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 LDAPEEMRYKACVTTTEAVASCVGVYFAMTKASWGGGKGKIRKVH-----NDDEV 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 86 VLYQVTRTGSFSFVSRSKAVOHGKPIFCQASFOQAPSPHQHOFSMPTVPPPELID 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 341 ALF-----KQVQGEVPGSPIFYMRLASQS-----RH-----LE 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 CETLIDQ-----LRDPLNOKRYPLALNRIAAQEVPIETKPVNPSLSQLQRMPEKPM 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 VOLLCDDQYGNVAALHSRDCSVQRH-----QKI-IEEGPVTVAPRETVKALEQAR 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 FVVRARYIGEGDMKHCVAAYISDYAFGLTALLPH-OMQHKV-----HPMV 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 RLAQAVGVGAATVEY--LYSMETGDYFVL--ELNPLQVEHPVTETAVNLPAAQAV 474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 246 SLDSHSMFHAPFRADHMLVCESPWAGGSRGLVHRLMRQDCVLA 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 475 GMGIPLQWQPEIRRFYGMDF-----GGGYD-----IWRKTAALA 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

```

US-08-695-421-6
: Sequence 6, Application US/08695421
: Patent No. 6268550
: GENERAL INFORMATION:
: APPLICANT: Gengenbach, B. G.
: APPLICANT: Somers, D. A.
: APPLICANT: Wyse, D. L.
: APPLICANT: Gronwald, J. W.
: APPLICANT: Egli, M. A.
: APPLICANT: Lucz, S. M.
: TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
: STREET: P. O. Box 2938
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/695,421
: FILING DATE: 23-AUG-1996
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/417089
: FILING DATE: 05-APR-1995
: APPLICATION NUMBER: 08/014326
: FILING DATE: 05-FEB-1993
: APPLICATION NUMBER: 07/917462
: FILING DATE: 21-JUL-1992
: APPLICATION NUMBER: 07/538674
: FILING DATE: 18-JUN-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Woessner, Warren D
: REGISTRATION NUMBER: 30,440
: REFERENCE/DOCKET NUMBER: 600.318052
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-373-6900
: TELEFAX: 612-339-3061
: TELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:

```

Search completed: August 19, 2002, 13:18:53
Job time: 24 sec

